

Resum

L'electricitat és un bé que no es pot emmagatzemar a gran escala. És a dir, la producció i el consum ha de coincidir en tot moment (descomptant les pèrdues per efecte Joule). Aquesta característica és bàsica per entendre el funcionament del mercat atès que el preu horari de l'electricitat del mercat majorista (tant pel mercat diari com per qualsevol dels mercats intradiaris) depèn en tot moment de l'oferta i la demanda en cada horitzó temporal.

HolaLuz.com, comercialitzadora elèctrica per la qual es realitza aquest projecte, actualment compra tota l'energia elèctrica en el mercat diari. Posteriorment ven aquesta electricitat en el mercat minorista i hi aplica un marge brut de 2 €/MWh en el cas de tarifes domèstiques i de 5 €/MWh en el cas de tarifes per PIME. Si es ponderen aquests marges per la cartera de clients, resulta un marge global de 2,5 €/MWh. HolaLuz.com treballa doncs amb un marge brut equivalent al 2,5 % de les vendes. És un marge estret, que fa que el negoci requereixi un gran volum per a ser rendible.

En el mercat diari es gestiona la major part de l'energia. De fet, el 71 % de l'energia generada en el mercat ibèric durant el 2013 s'ha negociat en aquest mercat. Tot i així, existeixen alternatives que podrien minimitzar el cost d'adquisició com poden ser els mercats intradiaris. Per tant, conèixer el comportament del preu de l'electricitat a curt termini pot suposar un gran avantatge competitiu atès que permetria escollir la millor opció de compra i repercutiria directament en el marge.

L'objectiu d'aquest projecte és dotar a l'empresa d'una eina per predir la diferència del preu horari de l'electricitat entre el mercat diari i l'intradiari 1. D'aquesta manera es podrà avaluar la diferència de preu en cada franja horària i escollir el mercat que minimitzi el cost.

Per tal d'estudiar el comportament dels mercats, es dissenya una metodologia basada en les sèries temporals amb introducció de variables exògenes i tractament de valors atípics. És a dir, models que tenen en compte tant valors passats del propi preu de l'electricitat com variables externes a aquest (demanda, energia eòlica, etc.). Aquests tenen com a finalitat predir la diferència del preu horari de l'electricitat i s'actualitzen al llarg de l'any 2012, permetent estudiar-ne l'evolució. Amb aquestes prediccions, s'aconsegueixen desviacions anuals d'un 12,99 % en el mercat diari i d'un 13,14 % en el cas de la primera sessió del mercat intradiari.

També es crea un model basat en xarxes neuronals per predir el mercat diari. L'objectiu és determinar si un model d'aquestes característiques podria utilitzar-se, amb garanties, per predir el preu de l'energia elèctrica. S'observa com, per a la predicció dels preus de l'electricitat, els models de sèries temporals s'adapten millor tant per la seva claredat com pels resultats obtinguts.

Finalment s'ha obtingut una eina que permet, a l'empresa, prendre decisions de negoci que impliquen millores en el seu marge brut d'entre el 2 % i el 17 % en funció de l'escenari que es contempla.

Índex

Resum	1
Índex de figures	8
Índex de taules	9
Glossari	11
Introducció	15
1 Definició del projecte	17
1.1 Objectius del projecte	17
1.2 Abast	17
1.3 Marc legal	18
2 El mercat d'energia elèctrica	21
2.1 El mercat majorista	21
2.2 Els mercats d'energia	22
2.2.1 El mercat diari	22
2.2.2 El mercat intradiari	22
2.3 Els processos d'operació tècnica	22
2.4 Funcionament dels mercats	23
2.4.1 Eix temporal	24
2.5 Factors que condicionen els mercats	24
3 Selecció del model	27
3.1 Estat de l'art	27
3.2 Models seleccionats	29
3.3 Software utilitzat	29

4	Model de sèries temporals	31
4.1	Conceptes previs	31
4.1.1	Sèries temporals i processos estocàstics	31
4.1.2	Estacionarietat i ergodicitat	31
4.1.3	Procés ARMA	32
4.1.4	Processos no estacionaris.	33
4.1.5	L'ACF teòrica	35
4.1.6	La PACF teòrica	35
4.2	Metodologia utilitzada per generar els models	35
4.2.1	Anàlisi univariant de sèries temporals	36
4.2.2	Anàlisi univariant amb variables exògenes	39
4.2.3	Determinació del model amb variables exògenes.	41
4.2.4	Anàlisi de valors atípics i linealització de la sèrie	41
4.2.5	Validació del model final	43
4.3	Selecció del mercat	45
5	Model de xarxes neuronals	47
5.1	Conceptes previs	47
5.1.1	Descripció d'una neurona artificial: El perceptró	47
5.1.2	Xarxa neuronal: El perceptró multicapa	48
5.1.3	Entrenament i validació d'una xarxa neuronal	49
5.1.4	Criteris de convergència	50
5.2	Configuració escollida	50
6	Resultats	53
6.1	Sèries temporals	55
6.2	Xarxes neuronals	65
6.3	Selecció del mercat de compra	66
7	Estudi econòmic	69
7.1	Desenvolupament de l'eina	69
7.2	Anàlisi de la inversió	70
8	Impacte ambiental i social	73
	Conclusions i futures línies de treball	76
	Bibliografia	78

Annexos 81

A Sèries temporals: Cas pràctic 81

A.1	Determinació del model univariant	81
A.2	Determinació del model amb variables exògenes	88
A.3	Detecció de valors atípics i linealització de la sèrie	89
A.4	Determinació del model final	95

B Codi R 99

B.1	Sèries temporals	99
B.1.1	Codi general	99
B.1.2	Funcions auxiliars	103
B.2	Xarxes neuronals	106

Índex de figures

2.1	Horitzó de temps dels sis mercats intradiaris. Extret del portal web d'OMIE.	22
2.2	Resultat d'una subhasta diària. Extret del portal web d'OMIE.	23
2.3	Factors utilitzats en diversos estudis. Extret de [3].	25
4.1	Esquema de la generació dels models i les prediccions previstes.	36
4.2	Representació esquemàtica dels tipus de valors atípics.	43
4.3	Diagrama simplificat de la metodologia aplicada.	44
5.1	Esquema d'una neurona artificial.	47
5.2	Espai de solucions en funció del número de capes extret de [18].	48
5.3	Possible sobreentrenament d'una xarxa neuronal.	50
6.1	Histograma del preu de cada període del mercat diari. L'eix horitzontal representa el preu (per MWh) en euros i el vertical el número de vegades que s'ha repetit aquest preu.	55
6.2	ACF i PACF mostrals utilitzats per identificar el model M1 del mercat diari.	56
6.3	Resum de l'evolució dels models. De color verd el període de generació del model i de color blau el període que ha predit.	59
6.4	Representació del MAPEc obtingut per a cada model entre 01/02/2012 i 01/05/2012 (mercat diari).	62
6.5	Representació del MAPEc obtingut per a cada model entre 02/05/2012 i 01/08/2012 (mercat diari).	63
6.6	Producció diària d'energia per tecnologies del mes d'abril. Extret del lloc web d'OMIE.	64
6.7	Sèries de preus de generació. L'eix horitzontal representen els dies i el vertical el logaritme del preu més una unitat. De color negre els valors considerats atípics.	64
6.8	Models estimats amb les xarxes neuronals pel primer període.	65
6.9	MAPEc per obtingut per cada model entre 04/01/2012 i 02/03/2012	66
A.1	Representació gràfica dels preus que generen M2.	81

A.2	Representació gràfica del logaritme neperià de la sèrie.	82
A.3	Representació de la setmana 14 de la sèrie que genera M2	83
A.4	Sèrie original amb una diferència d'ordre 24.	83
A.5	ACF i PACF mostrals.	84
A.6	Model univariant.	85
A.7	Representació del comportament dels residus. S'ha de tenir en compte que s'ha transformat la sèrie mitjançant logaritmes	86
A.8	Histograma i diagrama Q-Q dels residus.	87
A.9	ACF i PACF dels residus.	87
A.10	ACF i PACF dels residus al quadrat.	88
A.11	Model multivariant.	88
A.12	Representació dels valors observats. En negre els que es consideren atípics. . .	89
A.13	Valors atípics.	91
A.14	Valors compresos entre les observacions 500 i 512.	92
A.15	Valors compresos entre les observacions 575 i 585.	93
A.16	Valors linealitzats compresos entre les observacions 500 i 512.	94
A.17	Representació de la sèrie original menys la linealitzada.	95
A.18	ACF i PACF mostrals de la sèrie linealitzada.	96
A.19	Model multivariant final.	96
A.20	Representació del comportament dels residus de la sèrie linealitzada.	97
A.21	Histograma i gràfica Q-Q dels residus.	98

Índex de taules

6.1	Mitjana del preu per a cada mercat i període del 2012.	54
6.2	Models obtinguts per al mercat diari.	58
6.3	Models obtinguts per a la sessió 1 del mercat intradiari.	60
6.4	Mitjana del MAPEc obtinguda per a cada període del mercat diari i per a cada estratègia.	61
6.5	Resultats obtinguts aplicant la metodologia en funció de les hores de participació.	67
7.1	Cost de la inversió.	70
7.2	Estalvi per a cada cas i any.	71
7.3	Anàlisi de la inversió per al primer cas.	71
7.4	Anàlisi de la inversió per al segon cas.	71
7.5	Anàlisi de la inversió per al tercer cas.	72

Glossari

Símbols

∇^d	Operador diferència d'ordre 1 aplicat d cops.
$\tilde{\omega}_t$	Procés centrat.
O_i	Sortida i d'una neurona o xarxa neuronal.
X_j	Variable d'entrada j d'una neurona o xarxa neuronal.
α_{kk}	Coefficient d'autocorrelació parcial mostral d'ordre k .
∇	Operador diferència d'ordre 1.
∇_s	Operador diferència d'ordre S .
ω_t	Sèrie diferenciada.
$\bar{\omega}$	Mitjana de la sèrie diferenciada.
Φ	Coefficient del procés autoregressiu estacional.
ϕ	Coefficient del procés autoregressiu.
$\phi_p(B)$	Polinomi de la part autoregressiva.
ϕ_{kk}	Coefficient d'autocorrelació parcial d'ordre k .
ρ_k	Coefficient d'autocorrelació dels residus del retard k .
Θ	Coefficient del procés de mitjana mòbil estacional.
θ	Coefficient del procés de mitjana mòbil.
θ'	Factor d'activació.
$\theta_q(B)$	Polinomi de la part de mitjana mòbil.
ξ_t	Residu de l'instant t del model de regressió múltiple.
$AR(p)$	Procés autoregressiu d'ordre p .
B	Operador retard.
B^p	Operador retard d'ordre p .
D	Dia que es vol predir.
$D - 1$	Dia abans del que es vol predir. Dia de la predicció.

I_t	Grau de correcció del valor atípic localitzat l'instant t .
L	Versemblança.
M	Model del mercat diari.
$MA(q)$	Procés de mitjana mòbil d'ordre q .
Mi	Model de la primera sessió del mercat intradiari.
r_k	Coeficient d'autocorrelació mostral d'ordre k .
S_i	Node i d'una neurona o xarxa neuronal.
W_{ij}	Pes assignat a l'entrada j que convergeix en el node S_i .
$x_t^{(j)}$	Observació de l'instant t de la variable exògena j .
Z_t	Valor en l'instant t de la família de variables aleatòries del preu. Procés teòric.
z_t	Valor del preu observat en l'instant t . Realització de Z_t .
z'_t	Resultat d'aplicar logaritmes a z_t .

Sigles

AIC	<i>Akaike information criterion.</i>
AIC_c	AIC corregit.
ACF	<i>AutoCorrelation Function.</i>
AR	<i>Autoregressive.</i>
MA	<i>Moving Average.</i>
$MAPE$	<i>Mean Absolute Percentage Error.</i>
$MAPE_c$	$MAPE$ corregit.
$PACF$	<i>Partial AutoCorrelation Function.</i>
AO	<i>Additive Outliers.</i>
$ARIMA$	<i>Autoregressive Integrated Moving Average.</i>
$ARIMAX$	$ARIMA$ amb variables exògenes.
$ARMA$	<i>Autoregressive Moving Average.</i>
$CNMC$	<i>Comisión Nacional de los Mercados y la Competencia.</i>
CSN	<i>Consejo de Seguridad Nuclear.</i>
$GARCH$	<i>Generalized Autoregressive Conditional Heteroskedastic.</i>
LS	<i>Level Shift.</i>
$MAGRAMA$	<i>Ministerio de Agricultura, Alimentación y Medio Ambiente.</i>
$MIBEL$	<i>Mercado Ibérico de Electricidad.</i>
$OMIE$	<i>Operador del Mercado Eléctrico Polo Español.</i>

OMIP *Operador del Mercado Eléctrico Polo Portugués.*

PBCD Programa Base de Cassació Definitiu.

PBCP Programa Base de Cassació Provisional.

PDBF Programa Diari Base de Funcionament.

PDV Programa Diari Viable.

PDVP Programa Diari Viable Provisional.

REE *Red Eléctrica España, Operador del Sistema.*

SAIH *Sistema Automático de Información Hidrológica.*

SARIMA *Seasonal Autoregressive Integrated Moving Average .*

SARIMAX ARIMAX amb component estacional.

TC *Temporary Change.*

Introducció

Una comercialitzadora elèctrica es dedica a comprar electricitat al mercat majorista i a vendre-la als seus clients al mercat minorista. Aquest és el cas d'HolaLuz.com, comercialitzadora elèctrica on es realitza aquest projecte.

L'empresa lloga, a preus regulats, les xarxes de transport i distribució i posteriorment ven l'electricitat als clients finals aplicant-hi un marge brut de 2,5 €/MWh. Tenint en compte que el preu final d'un client oscil·la entre els 100 €/MWh i els 130 €/MWh, es fa evident que minimitzar el preu de compra és una tasca clau per ser més competitiu.

Les comercialitzadores poden adquirir aquesta energia al mercat diari, a les sis sessions del mercat intradiari o mitjançant contractes bilaterals. A més, poden participar en el mercat de futurs per tal d'assegurar-se un preu de compra fix durant un període de temps. Per tant, existeixen múltiples possibilitats i totes elles amb un cost diferent. Aquest cost d'adquisició no es sap fins que s'ha tancat cada mercat de manera que a priori, no sap quin mercat serà el més favorable.

Per tant, disposar d'una eina que permeti decidir a quin mercat s'ha de nominar l'electricitat és un gran avantatge competitiu. Obtenir l'electricitat a un preu més baix permet, a HolaLuz.com, incrementar el marge comercial i/o preus finals més competitius. A més, aporta una major solidesa financera a l'empresa.

Actualment l'empresa compra tota l'electricitat en el mercat diari de manera que aquesta eina permet plantejar-se alternatives. Tanmateix, com que el volum d'energia que negocia és un percentatge ínfim del mercat, qualsevol decisió que HolaLuz.com prengui no influeix en el comportament d'aquest (mercat).

Atès que el preu de l'electricitat es caracteritza per una periodicitat en el temps, s'estudien diferents models predictius que es basen en valors històrics del preu i altres possibles variables. Per determinar aquestes variables s'analitza el funcionament del mercat per determinar els factors que influeixen en el procés de casació. D'aquesta manera no és necessari un anàlisi exploratori de les variables.

És per això que l'origen d'aquest projecte rau en la voluntat d'HolaLuz.com d'obtenir una eina que permeti escollir en quin mercat s'ha de comprar l'electricitat per minimitzar el cost de compra. Aquest fet ha coincidit amb la inquietud (per part de l'estudiant) per entendre, des d'un punt de vista professional, el funcionament del mercat elèctric. Per tant, amb la voluntat d'extendre capacitats i aprofitant la necessitat d'HolaLuz.com de disposar d'una eina per diversificar les seves nominacions, s'ha considerant interessant realitzar aquest estudi. En aquest, s'han utilitzat tècniques estadístiques avançades que no s'han vist durant els estudis i ha sigut tutoritzat per un professor expert en sèries temporals.

Capítol 1

Definició del projecte

1.1 Objectius del projecte

L'objectiu principal del projecte és crear una eina que permeti escollir en quin mercat s'ha de nominar l'energia per minimitzar-ne el cost. Per poder assolir aquest objectiu final, es defineixen un conjunt d'objectius parcials.

- Predir el preu de l'energia elèctrica del mercat diari. Aquesta predicció s'haurà de fer el dia $D - 1$, abans de les 12:00 hores, per les 24 hores del dia D .
- Predir el preu de l'energia elèctrica de la primera sessió del mercat intradiari. Aquesta predicció s'haurà de fer el dia $D - 1$, abans de les 12:00 hores, per les 24 hores del dia D .
- Comparació de les prediccions d'ambdós mercats per determinar si es possible reduir el cost de compra.
- Determinació d'un model de sèries temporals i disseny d'una metodologia per obtenir-lo.
- Determinació d'un model de xarxes neuronals i analitzar la seva idoneïtat comparant-lo amb el model de sèries temporals.
- El temps d'ajust dels models no ha de superar els 30 minuts (un cop ajustat, la predicció és pràcticament instantània).

1.2 Abast

L'abast d'aquest projecte és predir la diferència del preu de l'electricitat del mercat diari i del mercat intradiari (any 2012) abans que HolaLuz.com faci la nominació horària pel dia següent. Amb aquesta, es prendrà la decisió de nominar en un mercat o en l'altre. Atès que es nominarà tota l'energia en un dels dos mercats, aquesta decisió s'haurà de prendre abans de que s'executi el primer mercat (en aquest cas és el mercat diari).

Actualment les nominacions horàries pel dia D , es realitzen el dia $D - 1$ abans de les 12:00 hores. Tot i que els preus analitzats són de l'any 2012, s'utilitzarà l'horari actual per tal que sigui compatible amb estudis posteriors.

Tot i així, serà necessari un marge de temps de manera que les prediccions es realitzaran a les 11:00 hores i en el cas que s'hagi d'estimar algun model, aquesta estimació es farà a les 10:00 hores. Per tant, en qualsevol dels casos es podran fer les prediccions a les 11:00 hores. Es recorda que un dels objectius és que els models s'estimin en menys de 30 minuts (en aquest cas, dos models en menys d'una hora). Per tant, les variables necessàries per realitzar les prediccions hauran d'estar disponibles abans de les 10:00 hores.

Per fer aquestes prediccions es construirà un model basat en les sèries temporals. L'obtenció d'aquest es basarà en la metodologia Box-Jenkins amb introducció de variables exògenes i tractament de valors atípics. Posteriorment es compararan les prediccions per tal d'escollir el mercat de compra, on es nominarà el 100 % de l'energia d'aquella hora.

També es realitzarà un model basat en xarxes neuronals per predir alguns mesos de l'any 2012. La finalitat és estudiar la seva idoneïtat pel cas d'estudi del present projecte.

Les variables exògenes utilitzades per a determinar aquests models són factors que intervien en el mecanisme de determinació del preu. Per tant, en cap cas es realitzarà l'anàlisi exploratori d'aquestes.

Per últim, es partix de la hipòtesi que les accions que pugui prendre HolaLuz.com respecte a les seves nominacions no afectaran en el comportament global del mercat. Aquesta suposició és assumible si es té en compte el baix volum d'energia que gestiona (l'any 2012 era aproximadament un 0,024% dels 250 TWh utilitzats durant tot l'any). Aquesta suposició es manté en el temps ja que HolaLuz.com aspira a comercialitzar 600 GWh l'any 2017, fet que suposa menys del 0,25 % del total. A més, es suposarà que els mercats dels anys posteriors al 2012 es comportaran d'una manera similar a aquest any. El compliment d'aquestes dues últimes condicions es necessari per tal que la solució sigui vàlida.

1.3 Marc legal

En un entorn de constant canvi, és important tenir en compte el marc legal del sector elèctric.

Sector elèctric general

Ley 54/1997, de 27 de noviembre, del Sector Eléctrico.

Llei que transposava la directiva europea 96/CE, la qual liberalitzava el sector elèctric.

Real Decreto 485/2009, de 3 de abril, por el que se regula la puesta en marcha del suministro de último recurso en el sector de la energía eléctrica.

L'1 de juliol del 2009 es va eliminar la *Tarifa Integral* i als consumidors se'ls va donar la possibilitat d'acollir-se a una tarifa regulada pel Govern (Tarifa d'Últim Recurs, oferida per les Comercialitzadores d'Últim Recurs) o contractar l'electricitat a una comercialitzadora del mercat lliure.

Ley 24/2013, de 26 de diciembre, del Sector Eléctrico.

Llei que deroga la *Ley 54/1997*. Les Comercialitzadores d'Últim Recurs només poden oferir les tarifes regulades mentre que la resta s'anomenen Comercialitzadores de Referència.

Regles del mercat majorista

Resolución de 30 de junio de 1998, de la Secretaría de Estado de Energía y Recursos Minerales por la que se aprueban las Reglas de Funcionamiento del Mercado de Producción de Energía Eléctrica y el contrato de adhesión a dichas reglas.

En aquesta resolució se desenvolupen les primeres regles de funcionament del mercat majorista d'electricitat.

Resolución de 1 de agosto de 2013, de la Secretaría de Estado de Energía, por la que se aprueban las reglas de funcionamiento del mercado diario e intradiario de producción de energía eléctrica y el cambio de hora de cierre del mercado diario.

En aquesta resolució es va retardar dues hores l'hora de tancament del mercat diari per tal que coincidís amb altres mercats europeus. Aquest nou horari s'aplica des del dia 15 d'octubre del 2013. La nova hora de tancament del mercat diari és a les 12:00 hores.

Resolución de 9 de mayo de 2014, de la Secretaría de Estado de Energía, por la que se aprueban las Reglas de funcionamiento de los mercados diario e intradiario de producción de energía eléctrica y el contrato de adhesión a dichas reglas.

Aquesta és l'última resolució publicada pel que fa a les regles del mercat.

Capítol 2

El mercat d'energia elèctrica

Per tal d'entendre el context en el qual es desenvolupa aquest projecte, aquest capítol explica de forma breu el funcionament del mercat elèctric. Abans d'entrar en detall, és important destacar certes particularitats del sector elèctric [1]:

- L'energia elèctrica no es pot emmagatzemar de manera eficient en grans quantitats. Per tant, no és un valor líquid.
- Sempre s'ha de cobrir tota la demanda. En la gestió del sistema no es considera la possibilitat d'abaixar la càrrega, amb l'excepció testimonial dels contractes amb interrompibilitat d'unes poques grans indústries.
- Inelasticitat de la demanda.
- Incerteses tant en la producció com en la demanda.
- En el mercat espanyol el preu està acotat entre 0 €/MWh i 180 €/MWh.

A més, presenta una elevada volatilitat que en dificulta el seu estudi [2, p. 69].

2.1 El mercat majorista

En el mercat majorista es determina l'energia elèctrica que es produirà i s'utilitzarà, tant en preu com en quantitat. En aquest, els generadors actuen com a venedors i les comercialitzadores (i alguns consumidors qualificats) ho fan com a compradors.

En aquest mercat de producció elèctrica hi ha dues possibles formes de contractació:

- Mitjançant la competència directa entre els venedors i compradors d'electricitat a través dels mercats d'energia del *Mercado Ibérico de Electricidad* (MIBEL). Aquests són gestionats per l'*Operador del Mercado Ibérico Polo Español* (OMIE) i s'anomenen mercats *spot*.
- Mitjançant contractes bilaterals directes entre els generadors i els compradors. Les quanties econòmiques són pactades lliurement entre les dues parts, però la seva resolució ha de ser comunicada a l'OMIE i a *Red Eléctrica España* (REE), l'Operador del Sistema, ja que han de ser gestionades a nivell tècnic.

Actualment, el 95 % de l'energia produïda i utilitzada al sistema elèctric espanyol és negocia al mercat *spot* i no a través de contractes bilaterals.

2.2 Els mercats d'energia

Al MIBEL existeixen dos tipus de mercats *spot*: el mercat diari i el mercat intradiari, del qual se'n celebren 6 sessions.

2.2.1 El mercat diari

El mercat diari és el que concentra un major volum d'energia negociat, gestionant durant el 2013 el 71 % de l'energia utilitzada en el mercat ibèric. En aquest hi participen els agents del mercat i té lloc el dia $D - 1$ abans de les 12:00 hores. Per cada franja horària del dia D , els compradors (comercialitzadores, consumidors directes i comercialitzadores d'últim recurs) i els productors presenten les seves ofertes de compra i venda a l'OMIE, respectivament.

2.2.2 El mercat intradiari

Un cop finalitzat el mercat diari (i fins les 12:45 hores del dia D) es realitzen 6 sessions on els compradors i venedors poden realitzar les seves ofertes per ajustar els seus programes de producció o consum. En ser un mercat posterior al diari, els agents participants tenen unes millors previsions i per tant, poden corregir la primera estimació.

Els mercats només poden negociar hores futures, de manera que, a excepció de les dues primeres sessions, la resta negocia menys de 24 hores (com es pot veure a la Figura 2.1).

S'ha de tenir en compte que els productors es poden convertir en compradors d'aquests mercats i viceversa. Podria donar-se el cas que han venut en el mercat diari una energia que no es podrà produir de manera que en el mercat intradiari la recompren. D'aquesta manera eviten les penalitzacions per desviaments. Per tant, es tracta d'un mercat d'ajustos.

En aquest mercat s'ha gestionat, durant el 2013, un 16,67 % de l'energia del mercat diari.

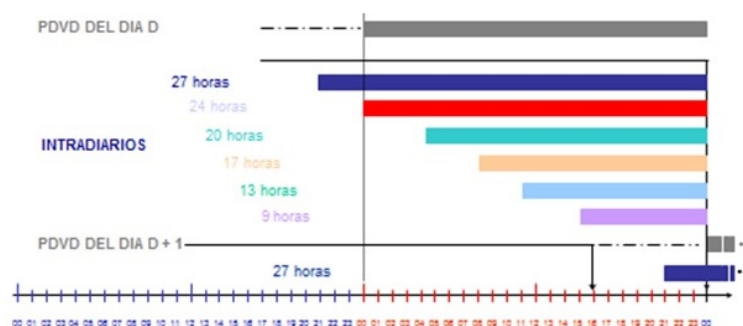


Figura 2.1: Horitzó de temps dels sis mercats intradiaris. Extret del portal web d'OMIE.

2.3 Els processos d'operació tècnica

Els resultats dels mercats són econòmicament viables, però s'ha de comprovar que també ho siguin tècnicament. Podria donar-se el cas que totes les unitats de producció que han resultat casades es trobin físicament a la mateixa zona geogràfica. Aquesta situació seria tècnicament inviable atès que la xarxa no podria suportar aquesta càrrega.

D'aquesta manera l'objectiu dels processos d'operació tècnica és garantir la seguretat i la fiabilitat del sistema de generació i transport d'electricitat. Aquestes gestions les porta a terme *Red Eléctrica España* (REE), l'Operador del Sistema, i deriva en uns costos de restriccions tècniques i serveis complementaris.

Les restriccions tècniques apareixen per les limitacions existents dins del propi sistema de transport d'electricitat. Els serveis complementaris són per garantir al sistema la capacitat per mantenir l'equilibri entre la generació i la demanda en les diferents fases (veure punt 2.4.1).

Per tant, del mercat no només se'n deriva el cost d'adquisició de l'energia elèctrica. En aquest estudi no es contemplaran els costos derivats dels processos d'operació tècnica, però es considera interessant la seva menció per tal d'orientar al lector.

2.4 Funcionament dels mercats

Tant el mercat diari com els intradiaris funcionen d'una manera similar. La principal diferència entre ells és la seva hora d'execució, la quantitat de franges horàries que es poden negociar i la seva liquiditat.

El mercat diari té lloc el dia $D - 1$ a les 12:00 hores de manera que els participants han d'enviar les seves propostes abans d'aquesta hora. Per a cada franja horària del dia D , els compradors presenten el preu màxim pel qual estan disposats a comprar una determinada quantitat d'energia mentre que els venedors indiquen el preu mínim pel qual estan disposats a vendre la quantitat d'energia horària que poden produir. Les ordres dels venedors poden ser complexes atès que hi pot haver alguna condició sobre el gradient de càrrega, ingressos mínims, condició d'indivisibilitat o parada programada. Per aquesta raó la casació pot ser simple (sense cap restricció) o complexa (amb condicions).

Les ofertes de venda són ordenades de menor a major preu mentre que les de compra s'ordenen en sentit invers. D'aquesta manera es determinen les corbes agregades de l'oferta i la demanda i es determina el preu de casació marginal. És a dir, el preu (marginal) ve determinat pel punt d'intersecció d'ambdues corbes per aquella franja horària. Queden fora de la subhasta tota oferta de venda amb un preu superior al de casació i tota oferta de compra amb un preu inferior a aquest. Aquestes serien totes les que queden a la dreta del punt d'intersecció de la Figura 2.2.

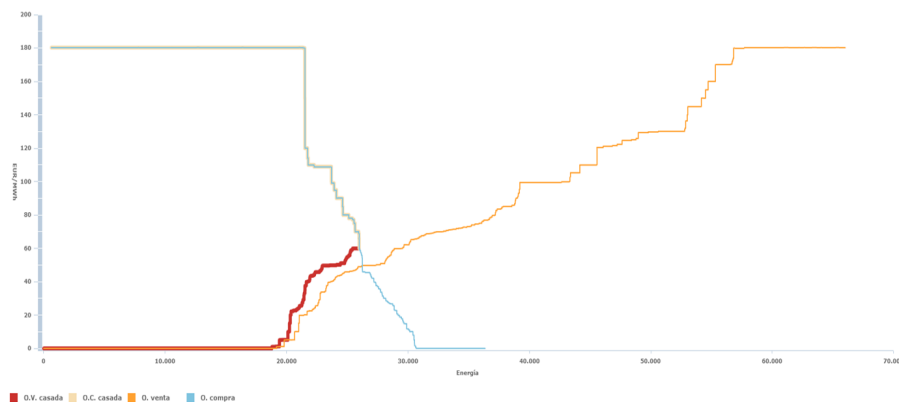


Figura 2.2: Resultat d'una subhasta diària. Extret del portal web d'OMIE.

2.4.1 Eix temporal

A continuació s'esmenta la seqüència d'operacions que afecten directament a la cassació del preu de l'electricitat del mercat diari:

- A les 12:00 hores, l'Operador del Mercat (OMIE) tancarà el període de recepció i validació d'ofertes realitzades pels agents del mercat (tant de compra com de venda).
- A partir de les 12:00 hores, i un cop s'han realitzat els anàlisis de les reclamacions presentades als processos de validació d'ofertes, s'iniciarà el procés de cassació. Els agents del mercat poden reclamar els resultats fins a 5 minuts després del tancament del període de recepció d'ofertes.
- A continuació es validaran les ofertes i en el cas que alguna no superi aquesta fase, no es tindrà en compte en el procés de cassació. Un cop s'han fet les validacions, s'executarà la cassació.
- Abans de les 13:00 hores, l'OMIE posarà a disposició de l'Operador del Sistema (REE) i dels agents del mercat el programa base de cassació provisional (PBCP) de forma confidencial. Els agents del mercat disposaran de trenta minuts per realitzar reclamacions sobre aquest. Alhora, l'OMIE podrà plantejar incidències.
- Abans de les 14:00 hores l'OMIE comunicarà als agents del mercat l'existència d'alguna reclamació o incidència pendent de resolució. L'horari de la seqüència podria ser modificat si la situació ho requereix.
- Si no hi ha reclamacions o incidències en el període de temps establert, el PBCP esdevindrà a programa base de cassació definitiu (PBCD).
- Abans de les 14:00 hores, un cop realitzat el procés d'incorporació de les energies declarades com contractes bilaterals, REE comunicarà a l'OMIE el programa diari base de funcionament (PDBF).
- Abans de les 16:00 hores, REE posarà a disposició de l'OMIE el programa diari viable provisional (PDVP), que haurà solucionat les restriccions tècniques. Aquest últim comunicarà als agents del mercat aquesta informació.

S'ha de tenir en compte que el preu marginal que pagaran els agents del mercat és el cassat en el PDBF. Per tant, les diferències que hi puguin haver amb el programa que finalment es realitza es compensaran amb els costos d'operació tècnica esmentats en el punt 2.3.

No s'ha exposat l'eix temporal dels mercats intradiaris atès que és el mercat diari el que limita el temps de decisió. Això és així atès que un cop s'hagi nominat en el mercat diari, aquesta posició no es pot desfer. Per tant, s'ha de saber abans de les 12:00 hores l'energia que es vol en cada mercat.

2.5 Factors que condicionen els mercats

Com s'ha vist, el preu de l'electricitat ve determinat per la intersecció de la corba de l'oferta i la demanda agregada. Així doncs, tota variable que afecti a aquestes corbes repercutirà sobre el mercat en el preu, la quantitat d'energia o en ambdós. Tot i així, la quantitat de factors que influeixen en el preu és tan elevada que determinar-los tots seria un projecte en si mateix. De fet, en diversos estudis s'han arribat a utilitzar fins a 40 factors diferents

[3]. A més, no tots ells són factors quantitativs, n'hi ha de qualitativs com són els factors psicològics o l'especulació.

A la Figura 2.3 es divideixen els factors que condicionen el mercat en 5 grans grups.

Tot i que la demanda és inelàstica, existeixen efectes temporals (C5). L'hora, el dia i el mes són factors determinants. A més, qualsevol factor que modifiqui els hàbits de l'ésser humà repercutirà en la demanda. Alhora, també hi ha una gran influència de la climatologia. Així, les temperatures extremes incrementen la demanda d'energia.

Pel que fa a la oferta, hi ha una gran varietat de factors que hi influeixen:

C1. Característiques del mercat. Són destacables les tecnologies utilitzades en la combinació energètica, però hi ha altres factors com els propis històrics de producció, intercanvis internacionals o històrics de les reserves energètiques.

C2. Incerteses no estratègiques. Aglutina les previsions de la generació i de les reserves, la temperatura, la climatologia i el preu de les matèries primeres.

Finalment, hi ha factors que afecten a les dues corbes. En aquest cas es troben els històrics de preus (C4) i altres incerteses estocàstiques com poden ser diferents índexs (C3).

Class	Input variable
C1	(1) Historical load (2) System load rate, (3) imports/exports, (4) capacity excess/shortfall (5) Historical reserves (6) Nuclear, (7) thermal, (8) hydro generation, (9) generation capacity, (10) net-tie flows, (11) MRR, (12) system's binding constraints, (13) line limits (14) Past MCQ (market-clearing quantity)
C2	(15) Forecast load (16) Forecast reserves, (17) temperature, (18) dew point temperature, (19) weather, (20) oil price, (21) gas price, (22) fuel price
C3	(23) Generation outages, (24) line status, (25) line contingency information, (26) congestion index
C4	(27) Historical prices
C5	(28) Demand elasticity, (29) bidding strategies, (30) spike existence index, (31) ID flag (32) Settlement period, (33) day type, (34) month, (35) holiday code, (36) Xmas code, (37) clock change, (38) season, (39) summer index, (40) winter index

Figura 2.3: Factors utilitzats en diversos estudis. Extret de [3].

Capítol 3

Selecció del model

Depenent del tipus d'informació de la qual es disposa, es convenient utilitzar un determinat tipus de model. Per aquesta raó es farà un petit anàlisi i es seleccionarà aquell que s'ajusti més a les necessitats de l'estudi. En aquest cas, es vol comparar la predicció de dos mercats diferents. Per aquesta raó es busquen models que realitzen prediccions puntuals del preu sense tenir en compte les incerteses.

3.1 Estat de l'art

Atès que l'energia elèctrica és un bé essencial per la societat, en els darrers anys s'han fet múltiples estudis amb l'objectiu de predir-ne el seu comportament. Atenent a les característiques dels models utilitzats per fer les prediccions, aquests es poden dividir en 4 grans grups.

En primer lloc hi ha els models estructurals, que intenten emular les relacions econòmiques i físiques del sistema per obtenir-ne la solució (és a dir, el preu). En segon lloc hi ha els models de teoria de jocs, que busquen l'equilibri entre els agents del mercat i el sistema. En tercer lloc hi ha els models de sèries temporals, que estudien sèries passades per estimar els coeficients dels models i predir valors. Finalment hi ha els models de xarxes neuronals, on mitjançant uns ajustaments dels pesos de les variables d'entrada, permeten obtenir prediccions. Tot i així, alguns analistes inclouen les xarxes neuronals dintre de les sèries temporals atès que per aquesta aplicació en concret hi ha moltes similituds [4].

Models de simulació o estructurals

Aquesta tipologia de model consisteix en crear un model que incorpori les relacions físiques i econòmiques del sistema elèctric. És a dir, es busca la solució utilitzant algoritmes regits pels fenòmens físics del procés i les seves conseqüències. Aquests algoritmes es modelen amb tècniques estadístiques i/o d'econometria. S'ha de tenir en compte que per realitzar un model d'aquesta tipologia és necessari disposar d'informació detallada del sistema segon a segon per poder predir amb exactitud el comportament del mercat. A més, és necessita una capacitat de processament molt elevada per tractar la informació [1].

Models d'equilibri o de teoria de jocs

Aquests models que identifiquen els agents participants en el mercat elèctric i busquen les decisions òptimes que han de prendre per maximitzar-ne el benefici. Aquestes decisions poden ser estratègiques, és a dir, cada agent participa tenint en compte les decisions de la resta. Modelitzar aquestes decisions és la part més complexa del procés.

Aquests models són especialment útils en un mercat on es coneixen els costos de producció i la concentració del mercat sense la necessitat de disposar d'un històric elevat de preus. [5, p. 135-159]

Dels models basats en aquesta metodologia en destaquen dos. El Model de Cost-Producció i el Model de Cost-Producció Estratègic. La diferència entre ells és que el primer ignora les decisions dels agents participants mentre que el segon les té en compte.

Ignorar la decisió estratègica pot resultar útil per mercats regulats on la incertesa és petita però és poc viable en cap altre cas.

Models de sèries temporals

La característica fonamental de les sèries temporals és que les observacions successives no són independents entre si de manera que l'anàlisi ha de tenir en compte l'ordre temporal de les observacions (mentre que els models estadístics tradicionals es basen en la independència de les observacions). Així doncs, les observacions passades de la variable resposta són un aspecte clau de l'anàlisi i per tant, s'ha d'estimar el seu efecte. [6, p. 101].

Les sèries temporals es poden classificar en dos tipus. Si els valors futurs es poden predir de forma exacta es tracta de sèries deterministes. En canvi, si es considera que tenen una distribució de probabilitat condicionada pels valors passats, es consideren sèries estocàstiques. Alhora, els models estocàstics poden ser univariants (només tenen en compte el propi comportament de la variable resposta) o multivariants (models univariants que incorporen variables exògenes).

Els models univariants més utilitzats són l'*ARMA* (*Autoregressive Moving Average*), i el *GARCH* (*Generalized Autoregressive Conditional Heteroskedastic*). El segon s'utilitza quan es vol complementar la predicció amb un interval de confiança que tingui en compte la volatilitat del moment.

Si altres variables afecten a la resposta es pot tenir present el seu efecte utilitzant models que incorporen altres variables. D'aquesta manera s'obtenen models univariants amb variables exògenes com l'*ARMAX* (*ARMA* amb variables exògenes).

Models de xarxes neuronals

Una xarxa neuronal està modelitzada seguint l'esquema del sistema nerviós d'un ésser viu. Per fer-ho, cada xarxa està formada per un conjunt d'unitats de processament que s'interconnecten convenientment. Aquestes unitats bàsiques s'anomenen neurones.

Cada neurona rep com entrada un conjunt de senyals (discretes o contínues) que processa i posteriorment transmet a les neurones que té connectades. Cada connexió té associada un pes que s'anirà modificant (de forma iterativa) segons una funció objectiu que minimitza l'error. D'aquesta manera s'arriba a l'estructura final amb la qual es podran realitzar prediccions.

Bàsicament existeixen dues configuracions de xarxa, la alimentada cap endavant (on la informació del grafs segueix una única direcció) i les que presenten cicles, anomenades recurrents. Aquests últimes són especialment útils per predir sèries temporals com podria ser el preu de l'electricitat, però el cost computacional és més elevat [7].

Un dels problemes de les xarxes neuronals és que tenen una topologia complexa. Per tant, no resulta evident la quantitat de neurones que han de formar-les i una bona elecció és fruit de l'experiència.

3.2 Models seleccionats

Atès que no es disposa d'informació privilegiada, els únics models viables són els basats en sèries temporals i en xarxes neuronals.

Pel que fa a les sèries temporals, són els models més utilitzats per la predicció de sèries que presenten periodicitat. A més, l'estudi de diversos processos teòrics permeten determinar les configuracions que millor s'adapten a les dades estudiades.

També resulta interessant realitzar l'estudi amb les xarxes neuronals però presenten certes dificultats. En primer lloc, tot i que en principi les xarxes recurrents són les que presenten millors resultats, representen un cost computacional molt elevat (i més si s'utilitzen variables exògenes). A més, l'elecció de la configuració interna de la xarxa resulta una tasca complicada.

Per tant, per la seva configuració senzilla, cost computacional relativament baix i idoneïtat per a sèries periòdiques, es decideix realitzar l'estudi mitjançant sèries temporals. Concretament s'utilitzaran models de tipologia *ARMA*.

Aquest es complementarà amb un petit anàlisi sobre la idoneïtat de les xarxes neuronals per la predicció d'aquesta tipologia de dades. En aquest cas s'utilitzarà una configuració cap endavant atès que les xarxes recurrents requereixen una capacitat de processament molt elevada.

3.3 Software utilitzat

Tots els estudis es realitzaran mitjançant el software estadístic *R*. Tot i que es coneixen altres softwares (com per exemple *MINITAB*® o *MATLAB*®), *R* presenta certes particularitats que fan que sigui dels més adequats per aquest projecte:

- Es tracta d'un software completament orientat a l'estadística i amb una gran versatilitat.
- És un software lliure. Per tant, el cost és nul i existeix una comunitat en continu creixement
- Llenguatge senzill i fàcil d'aprendre.
- Existència d'una gran varietat de funcions primitives que faciliten i agilitzen els estudis.
- Existència de paquets desenvolupats per investigadors que implementen metodologies avançades.
- Es un llenguatge suficientment potent malgrat no ser un dels més eficients.

Capítol 4

Model de sèries temporals

4.1 Conceptes previs

Abans d'explicar detalladament el procediment que es segueix per determinar els models de sèries temporals, es defineixen alguns conceptes. La justificació teòrica així com la nomenclatura utilitzada s'ha extret de *Análisis de series temporales* de Daniel Peña [8]. A més, s'han complementat alguns conceptes amb el material utilitzat en el Màster Interuniversitari en Estadística i Investigació Operativa de la UPC-UB (Universitat Politècnica de Catalunya i Universitat de Barcelona) [9].

4.1.1 Sèries temporals i processos estocàstics

S'anomena sèrie temporal a una seqüència ordenada d'observacions d'una mateixa variable resposta, on per cada instant t només hi ha una sola observació z_t . En molts casos, l'ordre de les observacions és important atès que hi pot haver dependència entre elles (com és el cas del preu de l'energia elèctrica).

Els valors observats, z_t , es poden considerar la realització d'una família de variables aleatòries Z_t . Aquest procés teòric, definit en el corresponent espai de probabilitat, es coneix com procés estocàstic.

4.1.2 Estacionarietat i ergodicitat

A continuació es defineixen els conceptes d'estacionarietat i d'ergodicitat atès que són condicions necessàries per poder determinar un model basat en els processos ARMA.

Una sèrie temporal és estacionària si la seva mitjana i la variància es mantenen constants per qualsevol instant de temps t i la funció d'autocovariància entre dues observacions només depèn del retard entre elles.

Una sèrie és ergòdica si l'autocovariància tendeixi a 0 quan el retard tendeix a l'infinit. És a dir, l'efecte del passat disminueix en augmentar el retard. Això implica que augmentar la mida mostral no aporta nova informació (sempre i quan es disposi d'una mínima quantitat de dades).

4.1.3 Procés ARMA

La manera més senzilla de relacionar linealment z_t amb z_{t-1} és mitjançant un procés autoregressiu ($AR(p)$):

$$z_t = \phi z_{t-1} + a_t \quad (4.1)$$

On ϕ és una constant ($|\phi| \leq 1$) i a_t (l'error) soroll blanc. El soroll blanc es defineix com una seqüència de variables que presenten una distribució normal amb mitjana nul·la.

La quantitats de retards que s'inclouen respecte z_t es coneix com l'ordre del procés (p). En l'equació 4.1 es pot observar com es tracta d'un procés autoregressiu d'ordre 1, $AR(1)$.

Generalitzant l'expressió per qualsevol valor de p s'obté:

$$z_t = \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + a_t \quad (4.2)$$

Per treballar més còmodament amb les dades, es defineix l'operador retard, B , com:

$$B^k z_t = z_{t-k} \quad (4.3)$$

Aplicant l'operador a l'equació 4.2 s'obté:

$$z_t = (\phi_1 B + \dots + \phi_p B^p) z_t + a_t \quad (4.4)$$

És a dir:

$$(1 - \phi_1 B - \dots - \phi_p B^p) z_t = a_t \quad (4.5)$$

Aquesta última expressió es pot escriure de forma compacte definint el polinomi característic $\phi_p(B)$:

$$\phi_p(B) = 1 - \phi_1 B - \dots - \phi_p B^p \quad (4.6)$$

Que substituint en l'equació 4.5:

$$\phi_p(B) z_t = a_t \quad (4.7)$$

També es pot buscar la relació entre les observacions tenint en compte les pertorbacions (o errors) anteriors. En aquest cas es coneix com a procés de mitjana mòbil (MA , de l'anglès *Moving Average*):

$$z_t = a_t + \theta a_{t-1} \quad (4.8)$$

On novament θ és una constant ($|\theta| \leq 1$) i a_t soroll blanc. L'ordre del procés es coneix com q i representa la quantitat de retards que s'inclouen respecte a_t . En l'equació 4.8 es pot observar com es tracta d'un procés $MA(1)$.

Aplicant l'operador B en el cas particular d'un $MA(1)$:

$$z_t = (1 + \theta B)a_t \quad (4.9)$$

Per tant, si es generalitza per qualsevol valor de q :

$$z_t = (1 + \theta_1 B + \dots + \theta_q B^q)a_t \quad (4.10)$$

Que també es pot escriure de manera compacta com:

$$z_t = \theta_q(B)a_t \quad (4.11)$$

Si s'inclouen ambdós processos en una única expressió s'obté un procés autoregressiu i de mitjana mòbil ($ARMA(p, q)$):

$$(1 - \phi_1 B - \dots - \phi_q B^q)Z_t = (1 + \theta_1 B - \dots + \theta_q B^q)a_t \quad (4.12)$$

Equivalent a:

$$\phi_p(B)Z_t = \theta_q(B)a_t \quad (4.13)$$

Amb $\phi \neq \theta$, la condició $|\phi| \leq 1$ garanteix que el procés sigui estacionari mentre que $|\theta| \leq 1$ garanteix que sigui invertible (equivalent a que el procés sigui ergòdic).

4.1.4 Processos no estacionaris.

Per poder determinar un model $ARMA(p, q)$ a partir d'una sèrie temporal és condició necessària que aquesta sigui estacionària. Per aquesta raó, tota sèrie que no presenti aquesta propietat haurà de ser transformada.

Variància no constant. Transformacions Box-Cox

Una de les possibles solucions per estabilitzar la variància d'una sèrie és aplicar transformacions Box-Cox $g(x)$. Aquesta família de transformacions potencials és la més utilitzada per estabilitzar sèries amb una variància creixent al llarg de les observacions:

$$g(x) = \begin{cases} \frac{x^\lambda - 1}{\lambda} & \text{si } \lambda \neq 0 \\ \log x & \text{si } \lambda = 0 \end{cases} \quad (4.14)$$

El valor de λ s'escull en funció de cada cas i x és la sèrie original.

Mitjana no constant. Diferenciació

Quan una sèrie presenta una tendència o estacionalitat (entesa com periodicitat en el temps), la mitjana no és constant. Tot i així existeix una transformació que elimina aquest efecte, la diferenciació. Es defineix l'operador diferència ∇ com:

$$\nabla = (1 - B) \quad (4.15)$$

De manera que:

$$\nabla z_t = (1 - B)z_t = z_t - z_{t-1} \quad (4.16)$$

Amb aquest operador es poden eliminar tendències de la sèrie original i s'anomena diferència de primer ordre. La quantitat de diferències aplicades d es denota com ∇^d .

També hi ha sèries que presenten estacionalitat. És a dir, existeix un patró de comportament que es repeteix cada S observacions. En aquest cas, s'han d'aplicar diferències d'ordre S per tal d'estabilitzar la mitjana:

$$\nabla_s z_t = (1 - B_S)z_t = z_t - z_{t-s} \quad (4.17)$$

S'ha de tenir en compte que cada cop que s'aplica una diferència es perden tantes observacions com l'ordre d'aquesta. Aquestes són les primeres de la sèrie. Per aquesta raó s'han d'aplicar les mínimes diferències.

Processos ARIMA i SARIMA

Després d'aplicar les diferències necessàries per tal que la sèrie temporal es comporti estacionàriament, aquesta es pot tractar com un procés ARMA.

Tot i així, existeixen processos que tenen en compte que cal diferenciar la sèrie original, facilitant-ne el seu maneig. Aquests incorporen una part integradora I al model original, donant lloc als processos ARIMA(p, d, q):

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - B)^d Z_t = (1 + \theta_1 B + \dots + \theta_q B^q) a_t \quad (4.18)$$

$$\phi_p(B) \nabla^d Z_t = \theta_q(B) a_t \quad (4.19)$$

És a dir, el propi procés s'encarrega de fer la diferència.

En el cas que la sèrie presenti estacionalitat d'ordre s , s'ha d'afegir (a la component regular) una component estacional:

$$(1 - \Phi_1 B - \dots - \Phi_P B^P)(1 - B^s)^D Z_t = (1 + \Theta_1 B + \dots + \Theta_Q B^{sQ}) a_t \quad (4.20)$$

Combinant aquesta part estacional amb la regular de l'equació 4.18 s'obté:

$$\begin{aligned} (1 - \phi_1 B - \dots - \phi_p B^p)(1 - \Phi_1 B - \dots - \Phi_P B^P)(1 - B)^d(1 - B^s)^D Z_t = \\ = (1 + \theta_1 B + \dots + \theta_q B^q)(1 + \Theta_1 B + \dots + \Theta_Q B^{sQ}) a_t \end{aligned} \quad (4.21)$$

Que amb notació compacta queda com:

$$\phi_p(B) \Phi_P(B^s) \nabla^d \nabla_s^D Z_t = \theta_q(B) \Theta_Q(B^s) a_t \quad (4.22)$$

Obtenint el procés SARIMA(p, d, q)(P, D, Q) $_s$.

4.1.5 L'ACF teòrica

L'autocorrelació simple d'ordre k ($k \geq 1$) d'un procés estacionari Z_t es defineix com:

$$\rho_k = \frac{Cov(Z_t, Z_{t+k})}{\sqrt{Var(Z_t)}\sqrt{Var(Z_{t+k})}} = \frac{\gamma_k}{\gamma_0} ; \rho_0 = 1 \quad (4.23)$$

Aquest és el coeficient de correlació lineal simple entre qualsevol parella de components de Z_t separats per un retard k .

Es defineix la funció d'autocorrelació simple (*ACF*, de l'anglès *AutoCorrelation Function*) del procés estacionari Z_t com la seqüència de valors que pren ρ_k .

4.1.6 La PACF teòrica

L'autocorrelació parcial d'ordre k ($k \geq 1$) d'un procés estacionari Z_t es defineix com el paràmetre ϕ_k de l'equació:

$$Z_t = \phi_{k1}Z_{t-1} + \phi_{k2}Z_{t-2} + \dots + \phi_{kk}Z_{t-k} + a_t \quad (4.24)$$

És a dir, és una mesura de l'associació lineal entre dos components de Z_t separats un retard k eliminant l'efecte dels retards intermedis i s'ajusta per mínims quadrats ordinaris.

Es defineix la funció d'autocorrelació parcial (*PACF*, de l'anglès *Partial AutoCorrelation Function*) del procés estacionari Z_t com la seqüència de valors que pren ϕ_{kk} ($\phi_{11} = \rho_1$).

4.2 Metodologia utilitzada per generar els models

El procediment per determinar el model (mercat diari) consta de tres fases. La primera, es basa en la metodologia Box-Jenkins [10], obtenint un model univariant. Posteriorment es complementa aquesta primera aproximació amb variables exògenes. Finalment, es fa un anàlisi dels valors atípics per tal d'obtenir el model final. Aquest anàlisi permet corregir aquests valors i per tant, obtenir un model més net.

Pel que fa al model per la primera sessió de l'intradiari, es fa exactament el mateix amb sèrie de preus corresponents. En aquest cas la diferència és que només s'utilitzen aquells models anàlegs als que s'han validat en el mercat diari. És a dir, si pel mercat diari s'utilitza el model $M1$ però no el $M2$, pel mercat intradiari es fa exactament el mateix (en aquest cas s'utilitzarà $Mi1$). Per tant, s'apliquen tots els punts menys el de la validació final, 4.2.5.

Aquesta última simplificació s'ha pres perquè el mercat diari i la primera sessió del mercat intradiari tenen un comportament similar.

Es considera que treballar amb finestres de 4 mesos aporta informació suficient per determinar un model adient atès que la previsió és a curt termini (a un dia). A la Figura 4.1 s'ha esquematitzat els períodes amb els quals es genera cada model i el seu horitzó de predicció previst. S'observa com el número del model correspon al mes que aquest està associat.

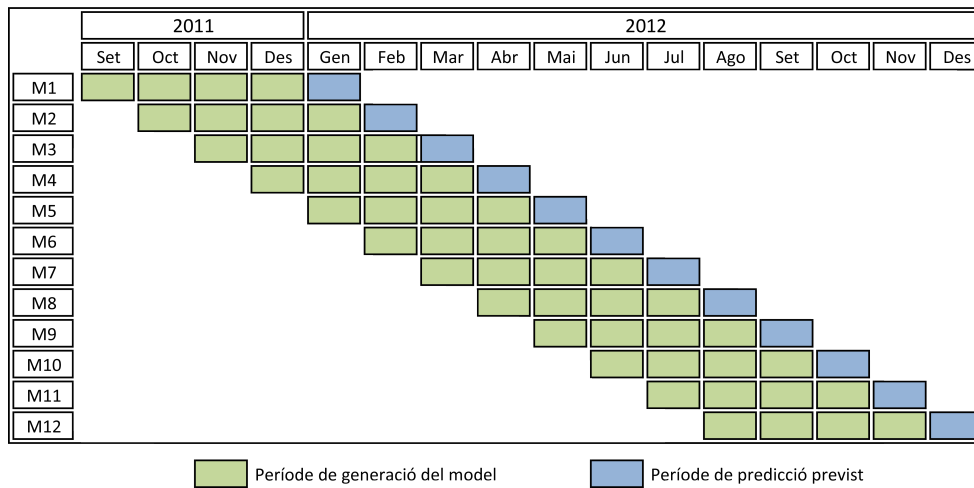


Figura 4.1: Esquema de la generació dels models i les prediccions previstes.

4.2.1 Anàlisi univariant de sèries temporals

Aquesta primera part de l'anàlisi es basa en la metodologia Box-Jenkins [10], metodologia utilitzada per determinar models ARMA. Aquesta consta de les següents parts:

- Identificació
- Estimació
- Validació
- Predicció

En el cas d'estudi no es fa la predicció atès que aquest model serveix per inicialitzar el model amb variables exògenes.

Identificació del model

En aquesta primera fase s'analitza la variable resposta, aplicant-li (si és necessari) les transformacions necessàries per tal que es comporti com una sèrie estacionària. Un cop s'han aplicat aquestes transformacions s'estudia el comportament de l'ACF i la PACF mostrals per identificar els possibles processos que les generen.

Petites variacions del preu en l'entorn de 0 €/MWh representen canvis percentuals molt grans. Per aquesta raó, abans de començar a tractar la sèrie, per corregir la mitjana i la variància, s'apliquen logaritmes:

$$z'_t = \ln(z_t + 1) \quad (4.25)$$

Per evitar qualsevol valor negatiu després de la transformació i problemes amb valors nuls, es suma una unitat a la sèrie original.

Transformació de la sèrie. Per poder determinar un model de la família ARMA(p,q) la sèrie ha de ser estacionària i ergòdica. D'aquesta manera primer es determina si és necessari aplicar transformacions per estabilitzar la variància i posteriorment si s'han d'aplicar diferències per fer el mateix amb la mitjana. L'ergodicitat de la sèrie es comprova en apartats posteriors.

És important tenir en compte que la diferència d'una sèrie estacionària continua essent estacionària. Per tant, s'han d'aplicar diferències el mínim número de vegades ja que es perdran $d + sD$ observacions.

1. Transformacions per estabilitzar la variància.
2. Determinació del número de diferències (d) d'ordre 1 necessàries.
3. Determinació del número de diferències (D) d'ordre s necessàries.

Cal notar que quan s'apliquen diferències a una sèrie a la que prèviament se li han aplicat logaritmes s'està treballant amb rendiments:

$$\nabla \ln z_t = \ln z_t - \ln z_{t-1} = \ln \left(1 + \frac{z_t - z_{t-1}}{z_{t-1}} \right) \approx \frac{z_t - z_{t-1}}{z_{t-1}} \quad (4.26)$$

Aquest concepte és molt utilitzat en el món financer atès que permet treballar amb increments relatius.

Anàlisi de l'ACF i la PACF mostrals. Sota condicions de normalitat, conèixer el procés teòric implica conèixer el vector de mitjanes i la seva matriu de covariàncies. Aquests elements poden ser inferits a partir de les observacions si es compleixen les condicions d'estacionarietat i d'ergodicitat. És a dir, si la sèrie (mostral) té un comportament semblant al d'un determinat procés (teòric), es podria suposar que la sèrie es comporta com aquell procés. Aprofitant aquesta propietat, la metodologia Box-Jenkins pretén identificar els paràmetres regulars (p, q) comparant les funcions d'autocorrelació mostrals amb les teòriques de diferents processos.

Pel que fa a la part estacional (P, Q), l'anàlisi és exactament igual que la part regular però només tenint en compte els valors de l'ACF i la PACF múltiples d' s .

Si ω_t és la sèrie diferenciada, els coeficients d'autocorrelació mostrals es calculen com:

$$r_k = \frac{\sum_{t=1+sD+1}^{n-k} (\omega_t - \bar{\omega})(\omega_{t+k} - \bar{\omega})}{\sum_{t=1+sD+1}^n (\omega_t - \bar{\omega})^2} \quad (4.27)$$

On k és el retard.

Els coeficients d'autocorrelació parcial α_{kk} s'obtenen de la següent equació:

$$\tilde{\omega} = \alpha_{k1}\tilde{\omega}_{t-1} + \dots + \alpha_{kk}\tilde{\omega}_{t-k} \quad (4.28)$$

On $\tilde{\omega}_t = \omega_t - \bar{\omega}$ és el procés centrat.

Les bandes de confiança tant per l'ACF i la PACF es situaran a $\pm 1,96/\sqrt{n}$. D'aquesta manera, tot retard fora d'aquesta banda es considerarà significativa.

S'ha de tenir present que cada procés teòric només té una funció d'autocorrelació associada, però una sèrie podria ser representada mitjançant diferents processos. D'aquesta manera és possible que s'identifiqui més d'un possible model. Si es presenta el cas, es prioritzen els models amb menys paràmetres, aplicant el principi de parsimònia.

Un cop trobades les funcions d'autocorrelació, es comprova que la sèrie presenta ergodicitat. Aquesta propietat es compleix si les funcions d'autocorrelació tendeixen a 0 per valors elevats de k .

Estimació del model

Un cop s'ha identificat el model més adient, se n'han d'estimar els seus coeficients. Per fer-ho, es força a la sèrie a comportar-se com el model identificat utilitzant el mètode de la màxima versemblança. Es poden consultar els detalls d'aquest càlcul en Box i Jenkins [10, p. 224].

Si s'han identificat diferents models amb la mateixa quantitat de paràmetres, el criteri que s'utilitza és el criteri d'informació d'Akaike (AIC , de l'anglès *Akaike information criterion*), escollint el de menor valor. Aquest dona una estimació relativa de la informació perduda d'un determinat model. És important tenir en compte que es tracta d'una paràmetre relatiu per comparar entre dos models i no aporta cap informació en sentit absolut:

$$AIC = 2K - 2 \ln L \quad (4.29)$$

On K és el número de paràmetres del model i L la versemblança.

Validació

En aquesta fase es comprova que les hipòtesis realitzades sobre els residus són certes. Aquests han de presentar mitjana nul·la, variància constant, distribució normal i no hi ha d'haver cap tipus de correlació entre ells. D'aquesta manera es comprova si el model estimat s'ajusta bé a les especificacions d'un model univariant estacionari. Si no es complissin aquestes condicions, s'hauria de tornar a identificar i estimar el model. Aquesta validació es fa per inspecció directa de diferents gràfics.

Abans de fer aquest anàlisi, la metodologia Box-Jenkins elimina del model tot coeficient no significatiu. En aquest cas no se n'elimina cap ja que algun podria passar a ser significatiu en introduir les variables exògenes.

Per validar el model, en primer lloc es representen els residus per comprovar que tinguin mitjana nul·la, variància constant i que siguin independents entre ells. En aquest cas es té en compte que fins un 5 % dels residus poden ser valors atípics. Després es representa l'histograma i la gràfica $Q - Q$ per comprovar que aquests tenen una distribució normal.

Acte seguit, per veure si existeix alguna dependència entre residus, es representen les seves funcions d'autocorrelació. En aquestes cap retard ha de ser significatiu (sempre s'ha d'assumir un 5 % de valors atípics que poden aparèixer com significatius).

Per últim es realitza el contrast de Ljung-Box per confirmar que no hi ha dependència entre residus. En aquest, la hipòtesi nul·la (H_0) és que els residus estan distribuïts independentment mentre que l'alternativa (H_a) indica una dependència.

$$Q = n(n+2) \sum_{k=1}^h \frac{r_k^2}{n-k} \quad (4.30)$$

On n és la mida mostral, h el retard que s'està avaluant i ρ_k el coeficient d'autocorrelació dels residus.

Q es distribueix com un khi quadrat amb $n - h - 1$ graus de llibertat.

Tot residu amb un p-valor per sota del 5 % implica dependència dels residus. Per tant, el model podria ser explicat introduint més coeficients al model original (és a dir, incrementar el valor de p o de q). En aquest cas s'ha de buscar una solució de compromís atès que incrementar indefinidament el número de retards del model no és una solució òptima.

Tenint en compte que es fan previsions a un dia i l'estacionalitat és de 24 observacions, seria ideal que com a mínim tots els p-valors fins a aproximadament l'error 24 fossin superiors al 5 %. De totes maneres, aquesta és una exigència molt elevada i tenint en compte que posteriorment s'afegeixen variables exògenes, es considera que si no hi ha dependència fins l'error número 10 serà suficient.

4.2.2 Anàlisi univariant amb variables exògenes

Com s'ha comentat, el mercat diari es pot veure afectat per molts factors. Per tant, per obtenir les prediccions, s'ha de tenir en compte aquestes variables i no només els valors històrics del preu.

Per tal de complementar l'anàlisi univariant, s'incorporen al model les variables exògenes que es consideren més significatives.

Amb aquestes variables es crea un model multivariant format per dues parts: una de regressió lineal múltiple i una altra de sèries temporals.

Selecció de les variables

Les variables utilitzades per fer els models s'han escollit tenint en compte diversos factors. Com s'ha comentat en el punt 2.4, es tracta d'un preu que prové del creuament de l'oferta i la demanda. Per tant, s'escullen variables que caracteritzin o condicionin aquestes dues corbes.

Previsió de la demanda horària. Tenir una bona aproximació de la corba de la demanda és un aspecte clau per obtenir un estudi sòlid.

En aquest cas es disposa de la previsió periòdica que fa REE. S'utilitza la previsió de la demanda pel dia D , disponible el dia $D - 1$. Aquesta està disponible en el portal d'informació de REE, e-sios [11], en l'apartat de previsions i està expressada en MWh.

S'ha de destacar que aquesta previsió és el millor estimador de la demanda fins el punt que s'adapta millor que el PDBF de la pròpia subhasta.

Previsió de l'energia eòlica horària. L'energia eòlica generada és una de les variables que més condicionen el preu de l'energia elèctrica atès que entra en el mercat diari a 0 €/MWh i amb un volum considerable. Així doncs, un percentatge elevat d'aquesta energia a la combinació energètica diària implica un preu més baix que si no fos així.

En aquest cas també es disposa d'una previsió de REE disponible el dia $D - 1$ a e-sios (MWh).

Potència nuclear diària. L'energia nuclear també entra a 0 €/MWh i amb un volum considerable.

En aquest cas no es disposa de cap previsió de manera que s'utilitzarà la potència generada en $D - 1$ a partir de les dades publicades pel *Consejo de Seguridad Nuclear* (CSN). És a dir, les dades disponibles just abans de fer la predicció. Per fer-ho, es seguirà el següent protocol:

1. El dia $D-1$ s'accedeix al lloc web del CSN a *Estados operativos de las centrales* [12].
2. S'accedeix als valors individuals de potència elèctrica generada de totes les centrals. S'escull el valor de l'última actualització disponible.
3. Es sumen totes les potències i s'assigna a cada hora del dia aquest valor. La potència generada per aquestes centrals és molt constant durant les 24 hores del dia de manera que és extrapolable a totes les hores. En aquest cas la potència està expressada en MW.

Reserves hidroelèctriques setmanals. Tot i que no representen directament la generació hidroelèctrica, quan les reserves augmenten, la probabilitat de que la generació també ho faci s'incrementarà [13, p. 21].

Tot i així, no deixa de ser una probabilitat. A més, una part de l'energia hidroelèctrica prové de centrals reversibles o de bombeig. La particularitat d'aquestes és que quan consideren que el preu de l'energia elèctrica es suficientment baix, acudeixen al mercat com consumidors i bombegen aigua per tenir-la disponible en moments on la producció sigui més rendible. És a dir, poden actuar com oferta a 0 €/MWh o com a demanda.

En aquest cas les dades s'extreuen del *Boletín Hidrológico*, publicat setmanalment (cada dimarts) pel *Ministerio de Agricultura, Alimentación y Medio Ambiente* (MAGRAMA) en seu portal oficial amb el nom de *Agua embalsada y energía disponible* [14]. Per tant, s'utilitza la dada disponible el dia $D - 1$ i es prendrà el mateix valor per les 168 observacions de la setmana.

Aquesta dada està expressada en GWh.

Preu diari del mercat de futurs. L'OMIP (*Operador Mercado Iberico Polo Portugués*) gestiona el mercat de derivats del MIBEL. Aquest és un mercat de futurs totalment financer. És a dir, no hi ha entrega física de l'energia contractada i es liquida per diferències amb el mercat diari. En aquest mercat es poden comprar valors futurs a un dia, a setmanes, a quadrimestres o a anys. S'utilitza com a variable el preu de tancament de l'OMIP dos dies abans del dia que es vol predir (repetit 24 observacions). Aquest és el valor disponible més proper al dia de la previsió i anterior a la subhasta i s'expressa en €/MWh [15].

Preu horari de la setmana abans. S'utilitza com a variable exògena el preu de l'electricitat de 7 dies abans. Aquesta no s'introdueix en l'anàlisi univariant per evitar que el model sigui excessivament gran (afegir una altra estacionalitat requereix molts recursos). A més, el software utilitzat no permet fer aquesta topologia de models. Aquest preu s'expressa en €/MWh i s'obté del web d'OMIE [16]. En el cas que s'hagi de generar l'històric complet, és recomanable descarregar els arxius del lloc web d'OMIP [15].

Per últim, quan s'introdueixen aquestes variables al model, es dividiran entre 10.000. Aquest fet ajuda a detectar la seva significació.

4.2.3 Determinació del model amb variables exògenes.

En primer lloc s'estima un model de regressió múltiple on la variable dependent serà el preu i les variables escollides les independents ($x_t^{(j)}$). Aquesta regressió és calculada mitjançant mínims quadrats:

$$z_t = \beta_1 x_t^{(1)} + \dots + \beta_j x_t^{(j)} + \xi_t \quad (4.31)$$

Reordenant els termes de la regressió s'obté:

$$\xi_t = z_t - \beta_1 x_t^{(1)} - \dots - \beta_j x_t^{(j)} \quad (4.32)$$

On la β_j és el coeficient de cada variable exògena j i ξ_t el residu.

És a dir, hi ha una component no explicada en el model univariant que ara si es pot explicar. Per aquesta raó es suposa que aquests residus tenen un comportament semblant al model univariant que s'ha identificat anteriorment.

D'aquesta manera, es tracta de crear un model ARMA per determinar el comportament de ξ_t .

$$\phi_p(B)\Phi_P(B^s)\nabla^d\nabla_s^D\xi_t = \theta_q(B)\Theta_Q(B^s)a_t \quad (4.33)$$

Obtenint el procés $SARIMAX(p, d, q)(P, D, Q)_s$.

Si es vol garantir el model òptim, aquesta estimació s'hauria de fer en una sola fase [17, pg. 355]. En aquest cas s'està estimant el model per fases degut a les limitacions del software. Tot i així, es considera que l'error que es pugui cometre és relativament baix.

A partir d'aquest punt es torna a aplicar la metodologia Box-Jenkins incrementant el nivell d'exigència. D'aquesta manera, se li exigeix al model que en el test de Ljung-Box mantingui independència dels residus fins al retard número 20.

Tot i així, tampoc es farà la predicció atès que abans s'haurà de fer un anàlisi dels valors atípics.

4.2.4 Anàlisi de valors atípics i linealització de la sèrie

L'existència de valors atípics a la sèrie de preus s'ha de tenir en compte atès que desvirtuen el model. Així doncs, aquests s'analitzen i si es convenient, es modifiquen creant una nova sèrie linealitzada. Amb aquesta nova sèrie s'estimen els nous coeficients, s'analitzen els residus i si es compleixen les condicions necessàries, es considera el model com a vàlid.

Per tal de determinar els valors atípics i la seva posterior linealització, s'utilitzen les rutines desenvolupades pel professorat de Sèries Temporals del Màster Interuniversitari en Estadística i Investigació Operativa de la UPC-UB. En B.1.2 s'adjunta el codi d'aquestes funcions, que reben el nom de *outdetect* i *lineal* respectivament.

Detecció de valors atípics

En primer lloc, s'ha d'establir el criteri per considerar un valor com a atípic. És a dir, s'han de determinar les bandes fora de les quals el valor és considerat com anòmal. Per situar-les s'utilitza la desviació estàndard com a mesura aproximada. Es tracta d'un valor aproximat perquè cada cop que es detecta un atípic a la sèrie la desviació estàndard es torna a estimar.

Allunyant molt les bandes, no es millora en excés el model ja que hi ha pocs valors atípics en unes sèries molt volàtils, mentre que si són molt estretes es determina un model ideal que no s'ajustarà a la realitat. Per tant, es necessari arribar a un equilibri.

El criteri per determinar si la col·locació de les bandes és adequada és que no més d'un 5 % de les dades representin valors atípics i alhora, una millora substancial de la variància del model.

És important tenir en compte els diferents tipus de valors atípics que hi pot haver atès que no tots es poden tractar de la mateixa manera (Figura 4.2).

- AO (*Additive Outliers*). Són observacions puntuals fora del valor. Per corregir-los només cal substituir-los pel valor esperat segons el criteri establert.
- LS (*Level Shift*). En aquest cas es tracta d'un conjunt d'observacions que han patit un canvi de nivell. És a dir, trenquen totalment amb la tendència i posteriorment no la segueixen sinó que en creen una de nova. Aquests casos són més conflictius atès que substituir els valors anòmals no és la solució. De fet, no són valors atípics per la nova tendència.
- TC (*Temporary Change*). Es tracta d'un LS però amb una recuperació posterior, tornant a la tendència original. Així doncs, es poden substituir els valors transitoris pels valors esperats.

Linealització de la sèrie

Un cop s'han determinat els valors atípics i s'ha comprovat que no s'ha comès cap error (per exemple per una mala introducció de les dades), aquests es poden corregir. En aquest cas, es corregiran els AO i els TC. Els LS són casos molt excepcionals on la millor solució és que el propi model identifiqui la nova tendència el més ràpid possible.

D'aquesta manera s'obté una nova sèrie de preus idèntica a l'original però amb els valors atípics substituïts pels valors que el model esperava trobar. La diferència entre aquests dos valors pel valor atípic i s'anomenarà ω_i :

$$\xi_{it} = z_t - \beta_1 x_t^{(1)} - \dots - \beta_j x_t^{(j)} - \sum_{i=1}^a \omega_i I_t \quad (4.34)$$

On a són els valors atípics i I_t el grau de correcció.

Si es corregeix un AO, I_t és 1, mentre que si s'està corregint un TC, el grau de correcció depèn de quants valors transitoris hi hagi. Aquest valor val $0,7^{(t'-t_0)}$ on t' és el valor on s'avalua i t_0 el valor on s'ha iniciat el TC.

Amb aquesta nova sèrie linealitzada es repeteix tot el procediment des de l'inici fins obtenir un model multivariant i fer-ne el corresponent anàlisi dels residus. En aquest cas el grau d'exigència és com el del punt 4.2.2, buscant una millora substancial de la variància.

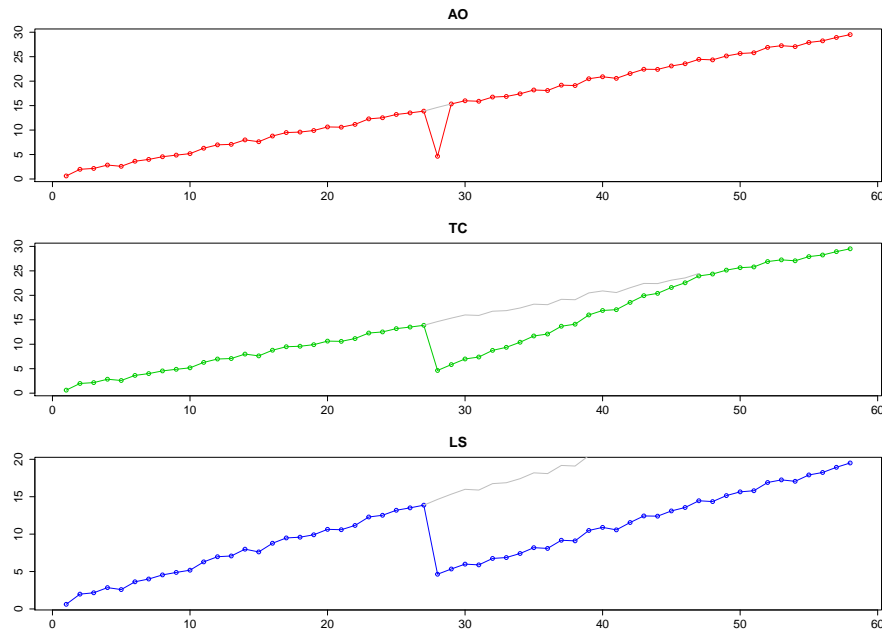


Figura 4.2: Representació esquemàtica dels tipus de valors atípics.

Si les bandes s'han col·locat correctament, el nou model té millors propietats que el model original i unes prediccions més fiables. Tot i així, si aquestes bandes són massa curtes, aparentment apareixerà un model molt bo amb una variància molt baixa, però realment no serà així ja que el model provindrà de dades ideals. Això es comprova a la validació del model.

Quan es treballa amb la sèrie linealitzada, el valor AIC s'ha de corregir atès que cada un dels valors atípics que s'han modificat afegeixen un paràmetre al model. De l'equació 4.29 es pot deduir que:

$$AIC_c = AIC + 2a \quad (4.35)$$

4.2.5 Validació del model final

Un cop s'ha comprovat que els residus es comporten correctament i es supera la prova de Ljung-Box, s'ha de validar el model. Es recorda que aquesta fase només es farà en el cas del model del mercat diari. Per tant, si s'accepta el model M_i (model mercat diari), s'acceptarà el model M_{i_i} (model mercat intradiari).

Generar un model mensual té per finalitat poder seguir possibles canvis de tendència, però també protegir als models de períodes de generació amb excessius valors atípics. D'aquesta manera si un model ha detectat el canvi de tendència és convenient actualitzar-lo mentre que si s'ha generat amb una sèrie desvirtuada, és preferible continuar amb el model anterior.

Per decidir si es manté o s'actualitza el model s'analitza la quantitat de valors atípics continguts en l'últim dels quatre mesos de generació i s'aplica una heurística.

En una situació ideal, la quantitat de valors atípics de cada període hauria de ser aproximadament un 25 % del total. Per tant, si aquest percentatge és més elevat, implicarà un comportament més anòmal. La cota esmentada es situa en el 40 %

S'utilitza aquest criteri per evitar validar amb uns dies després de la generació. Si aquests dies fossin valors atípics, el criteri de validació seria molt vulnerable. D'aquesta manera es pot validar tenint una millor perspectiva de la situació.

En el cas que sigui el primer model que es genera i es supera la cota de validació, aquest serà descartat.

La Figura 4.3 resumeix tota la metodologia que s'ha exposat.

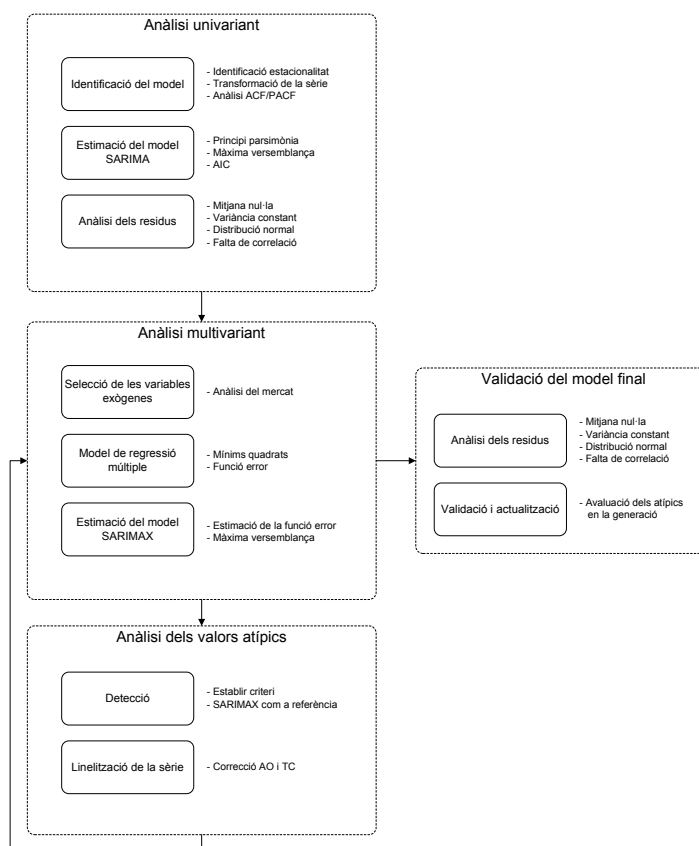


Figura 4.3: Diagrama simplificat de la metodologia aplicada.

Valoració de les prediccions

Un cop feta la predicció, el criteri per avaluar el model es basa en el valor de la mitjana dels errors relatius (*MAPE*, de l'anglès *Mean Absolute Percentage Error*):

$$MAPE = \frac{1}{24} \sum_{k=1}^{24} \frac{|\tilde{z}_t - z_t|}{z_t} \quad (4.36)$$

Tot i així, l'existència de situacions degudes a causes assignables no predictibles distorsionen aquesta mesura de valoració. Com es pot observar, una sola observació amb valor nul farà que el $MAPE$ sigui infinit i per tant, té sentit descartar-les. Per aquesta raó es treballa amb un $MAPE$ corregit ($MAPE_c$) on no es comptabilitza qualsevol error relatiu superior al 100 %.

4.3 Selecció del mercat

Atès que el mercat diari i la primera sessió del mercat intradiari presenten moltes similituds, es genera els models per la sessió 1 utilitzant les mateixes variables. Tot i així, si en un futur s'incorporen variables a un dels dos mercats, es podrà fer sense cap problema.

Així doncs, per cada hora, es comparen les prediccions obtingudes per cada mercat de manera que s'escull aquell que presenti una previsió del preu menor. En cas d'empat, s'escull el mercat diari, opció que utilitza l'empresa per defecte.

Capítol 5

Model de xarxes neuronals

5.1 Conceptes previs

Abans d'explicar detalladament el procediment que es segueix per determinar el model de xarxes neuronals, es defineixen alguns conceptes.

5.1.1 Descripció d'una neurona artificial: El perceptró

El primer model neuronal modern va ser proposat per McCulloch i Pitts al 1943. Aquest intentava imitar el comportament d'una neurona biològica realitzant una suma ponderada dels valors d'entrada. Posteriorment a aquesta suma se li afegia un factor d'activació (θ') i al conjunt se li aplicava una funció no lineal (per exemple, una funció graó).

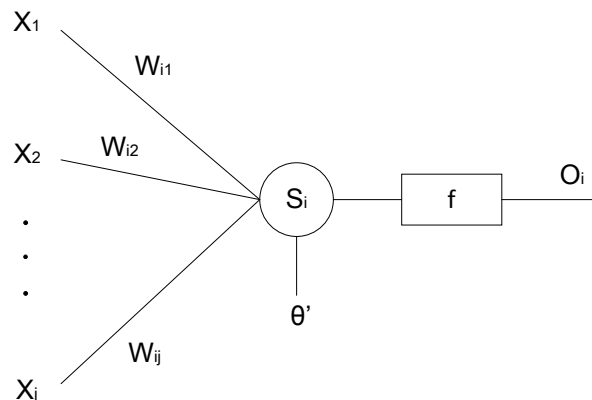


Figura 5.1: Esquema d'una neurona artificial.

A la figura 5.1 es pot observar la configuració d'una neurona on X_j són les entrades i W_{ij} els pesos de cada una d'aquestes entrades. Aquesta suma ponderada convergeix en un en

un node (S_i), se li suma un factor d'activació (θ') i posteriorment s'aplica una funció de transferència no lineal (f), obtenint la sortida O_i . Per tant, el valor de cada node serà:

$$S_i = \sum_j W_{ij} X_j + \theta' \quad (5.1)$$

El perceptró simple és una estructura com l'anterior on es compara el valor de sortida obtingut amb el valor desitjat. D'aquesta manera es pot calcular l'error i per tant, modificar el valor dels pesos per minimitzant-lo. Aquest procés iteratiu s'anomena entrenament supervisat:

- Inicialització dels pesos
- Determinació de la sortida O_i
- Càlcul de l'error
- Actualització del valor dels pesos

5.1.2 Xarxa neuronal: El perceptró multicapa

Una xarxa neuronal és un conjunt de neurones interconnectades convenientment. Si aquesta xarxa està formada per perceptrons simples (convenientment connectades) s'obté és el perceptró multicapa. La detecció de patrons en aquesta configuració es basa en la interconnexió paral·lela. D'aquesta manera s'aconsegueix una gran robustesa i es divideixen les diferents entrades en múltiples funcions lineals, detectant patrons complexos (Figura 5.2).

Cada conjunt de neurones connectades paral·lelament s'anomena capa. D'aquesta manera es té una capa d'entrada, una de sortida i es poden tenir diverses capes ocultes. Tot i que es poden construir xarxes amb diverses capes ocultes, es pot demostrar que la utilització d'una única capa oculta és suficient per tal que la xarxa actuï com un aproximador universal de funcions [18, pg. 52].

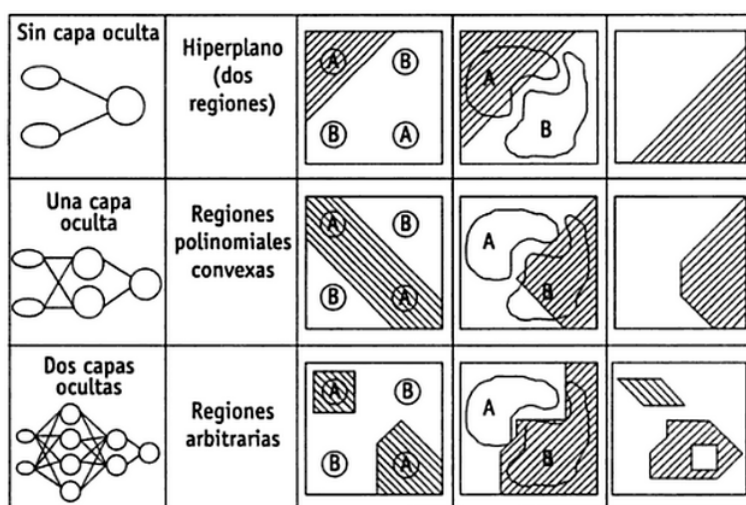


Figura 5.2: Espai de solucions en funció del número de capes extret de [18].

Aquesta configuració de xarxa que s'ha exposat s'anomena xarxa alimentada cap endavant (de l'anglès, *feedforward*). Es tracta d'una xarxa que no presenta cicles on la seva implementació i simulació és senzilla. Per tant, la informació es va transmetent en una sola direcció des dels nodes d'entrada fins als de sortida. Aquest fet permet una gran velocitat de processament.

També existeixen les xarxes neuronals recurrents. Aquestes presenten cicles i són molt utilitzades per la predicció de sèries temporals. Tot i així, s'ha de tenir en compte que la seva velocitat de processament és més baixa que el cas anterior. A més, si es disposa d'una gran quantitat de variables, és una opció poc viable.

Com es disposa d'una gran quantitat de variables s'ha optat per una xarxa alimentada cap endavant.

5.1.3 Entrenament i validació d'una xarxa neuronal

Bàsicament existeixen dos fases en la modelització d'una xarxa neuronal: l'entrenament i la validació.

L'entrenament utilitza un subconjunt de dades per determinar els pesos que defineixen la xarxa. És un procés iteratiu on s'intenta minimitzar l'error entre la sortida obtinguda per la xarxa i la sortida desitjada.

Existeixen diferents tipus d'entrenaments supervisats per les xarxes neuronals cap endavant. Un dels entrenaments que ofereix millors resultats és la retropropagació. Aquest consisteix en calcular l'error a la sortida i propagar-lo cap endarrere. D'aquesta manera, cada neurona es corregeix en funció de la seva contribució al resultat i aprèn a detectar uns patrons.

S'ha de tenir en compte que l'espai de solucions pot presentar mínims locals. És a dir, no es pot assegurar que l'entrenament hagi portat a la solució òptima. Per aquesta raó, si no s'assoleixen els objectius mínims, es pot optar per canviar la topologia de la xarxa, fer un altre entrenament amb uns pesos inicials diferents, modificar l'ordre de les dades d'aprenentatge o qualsevol altra acció que la modifiqui.

Si s'entrena excessivament una xarxa, aquesta pot patir sobrentrenament. Aquest fenomen té lloc quan la xarxa ha perdut la capacitat per generalitzar, adaptant-se molt bé al subconjunt d'entrenament i molt malament a la resta. Per aquesta raó l'entrenament ha de ser representatiu i generalitzat. A la Figura 5.3 s'observa una mostra que entrena la xarxa. Per obtenir una solució generalitzada, l'entrenament hauria de portar a la línia discontinua. En el cas que s'entreni excessivament s'arribarà a la línia continua, perdent la capacitat per predir altres mostres.

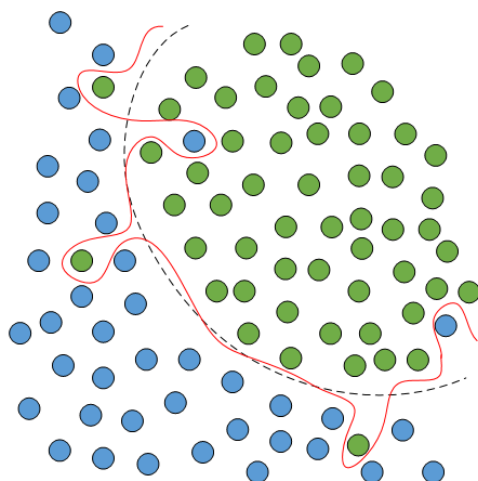


Figura 5.3: Possible sobreentrenament d'una xarxa neuronal.

Un cop la xarxa ha sigut entrenada, cal comprovar que es comporta adequadament. Per fer-ho s'utilitza el subconjunt de dades restant. Aquest procés de control posterior a l'entrenament s'anomena validació.

5.1.4 Criteris de convergència

És important evitar el sobreentrenament, però si una xarxa s'ha entrenat massa poc, també obtindrà uns mal resultats. Per aquesta raó s'han d'establir criteris per determinar quan ha de parar l'entrenament (i així també consumir els recursos justos i necessaris).

- Número de cicles: Es pot acotar el número d'iteracions de l'entrenament. D'aquesta manera es pot avaluar si una xarxa és massa complicada i requereix masses recursos.
- Nivell d'error: Es pot entrenar una xarxa fins que l'error comès (o la seva derivada) sigui inferior a una cota. Establir una cota massa elevada pot portar a un entrenament poc exhaustiu mentre que una cota massa baixa pot portar al sobreentrenament i/o una quantitat d'iteracions massa gran.
- Modificació ínfima del pes: Quan s'estan fent diverses iteracions i els valors dels pesos varien molt poc, es pot considerar que s'està molt a prop del mínim local. D'aquesta manera es pot establir una cota per detectar aquest fenomen.

5.2 Configuració escollida

Tot i que s'estudia una sèrie temporal, es decideix utilitzar una xarxa neuronal cap endavant ja que la quantitat de variables és elevada. Aquesta constarà d'una primera capa d'entrada formada per 99 nodes, una d'oculta de 24 i finalment una de sortida de 24 que representarà el preu per cada hora.

Pel que fa a les entrades, s'adequaran les variables del model de sèries temporals a aquest cas. D'aquesta manera s'han utilitzat 24 valors per cada variable de la que es disposa informació horària (previsió de la demanda, previsió de l'energia eòlica i el preu d'una setmana enrere)

i de la resta, s'ha utilitzat un únic valor. A més, s'ha afegit com entrada els 24 preus del dia immediatament anterior.

Per tant, s'està simplificant el model a un $AR(1)$ amb estacionalitat. Aquest és un dels inconvenients de les xarxes neuronals atès que s'està perdent la seqüencialitat.

L'algoritme d'entrenament que s'utilitzarà serà el de retropropagació (concretament el *PROP+*, disponible en el paquet *neuralnet*) utilitzant mínims quadrats.

Es realitzaran 20 repeticions i es limitarà cadascuna a $4 \cdot 10^5$ iteracions o quan la derivada de l'error sigui inferior a 0,01. D'aquestes 20 repeticions, s'escollirà el model amb un error menor.

Com les sèries temporals, es generarà el model amb una finestra de 4 mesos, reservant 15 dies (aleatoriament escollits) per fer la validació.

Capítol 6

Resultats

Els resultats es dividiran fonamentalment en dues parts. En la primera s'estudiarà el comportament dels models obtinguts, les prediccions i altres temes que puguin condicionar la qualitat de les prediccions. S'analitzarà detalladament el model diari atès que s'hauran generat els 12 models. Es recorda que dels models de la sessió 1 (de l'intradiari) només s'han generat els anàlegs als utilitzats pel mercat diari (simplificant tot el procediment).

Posteriorment s'analitzaran els resultats obtinguts de la comparació d'ambdós mercats. D'aquesta manera es podrà determinar si comparant aquests models es poden reduir els costos de compra.

Es considera que dividint els resultats d'aquesta manera s'aconseguirà, per una banda, analitzar econòmicament els resultats (objectiu final del projecte) i per l'altra, extreure resultats i conclusions que puguin millorar els models.

Anàlisi dels mercats

Abans d'analitzar els resultats obtinguts, és interessant tenir en compte el comportament del preu en el períodes predits. Conèixer com s'ha comportat el preu ajudarà a interpretar els resultats i per tant, guiarà en els aspectes que es consideren interessants estudiar.

En primer lloc s'analitzarà la mitjana del preu per cada període. Com es sap que el principal problema d'aquest tipus de prediccions són els valors atípics (en general valors mol baixos del preu), és possible que la mitjana indiqui en quins períodes hi pot haver un comportament no desitjat. A la Taula 6.1 s'han representat les mitjanes del preu de cada període que s'ha predit. Cada període correspon al mes que representa el seu número, tenint en compte que l'últim dia de generació sempre és un dimarts.

Període	Mercat diari [€/MWh]	Sessió 1 [€/MWh]
1	52,32	52,30
2	54,11	53,21
3	47,61	47,11
4	38,63	38,43
5	45,14	43,98
6	52,74	50,96
7	50,34	48,21
8	49,33	49,08
9	48,09	47,34
10	43,62	43,18
11	43,96	44,06
12	43,27	42,40

Taula 6.1: Mitjana del preu per a cada mercat i període del 2012.

S'observa que el preu de la sessió 1 és en quasi tots els períodes lleugerament inferior al del mercat diari. De fet, la mitjana del mercat diari és de 47,49 €/MWh mentre que el de la sessió 1 és 46,74 €/MWh. És possible que en aquest context econòmic, el terme de la cointegració s'adapti al comportament observat.

Tenint en compte que, en general, els valors atípics es troben en valors baixos del preu, és possible que els períodes amb un preu més baix que la mitjana presentin alguna particularitat. Aquests períodes són el 4, 5, 10, 11 i 12 pels dos mercats.

Tot i així, aquesta informació és massa pobre, raó per la qual el representen els histogrames de cadascun dels períodes amb l'objectiu de poder analitzar com es distribueixen els preus (Figura 6.1).

Tots els períodes presenten una distribució del preu asimètrica (negativa). A més, a excepció del període 5 (que té la mitjana més elevada), tots els períodes esmentats anteriorment presenten una particularitat. S'observa com la distribució és semblat però amb la freqüència dels valors mínims incrementada. És a dir, la baixada de la mitjana es deguda a aquests valors que possiblement són atípics.

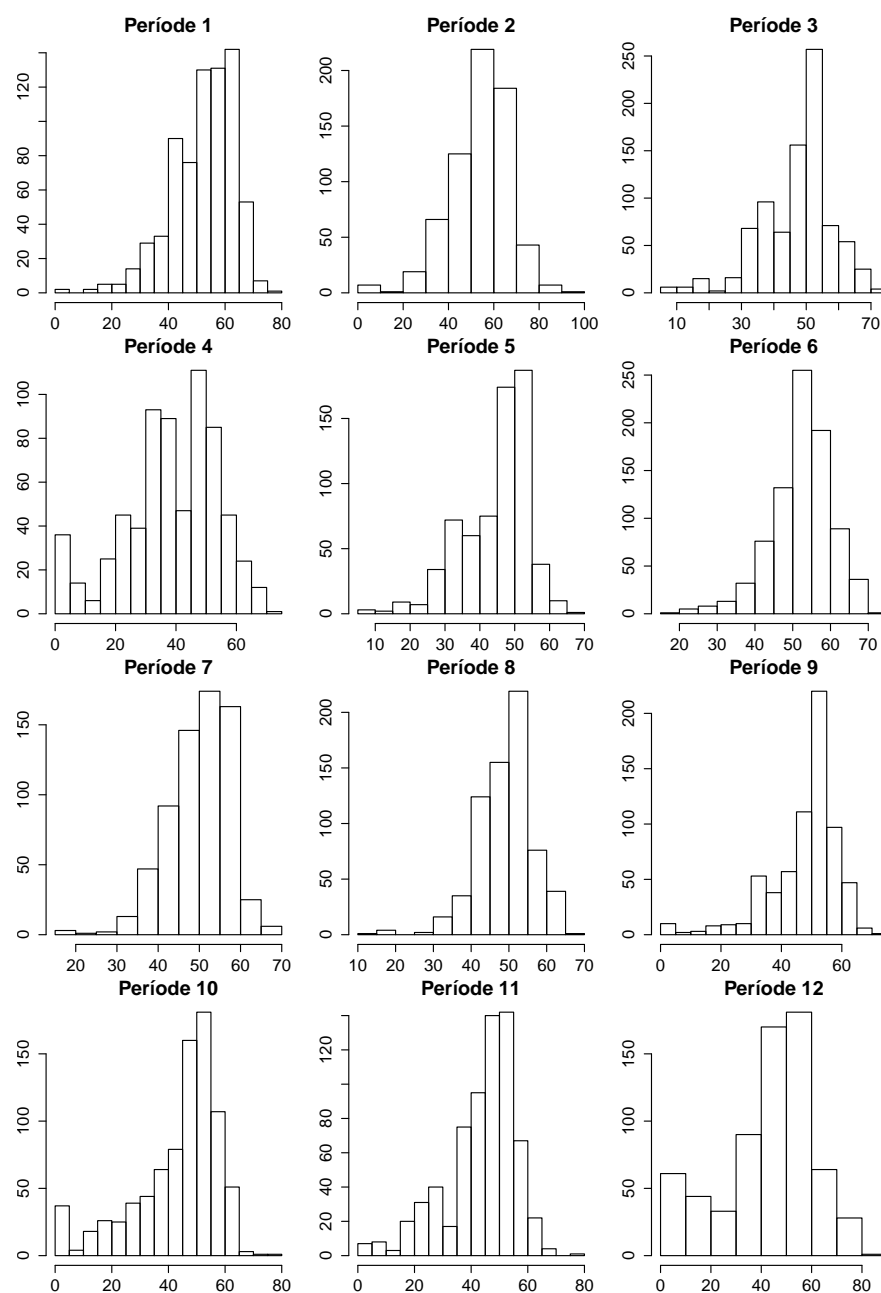


Figura 6.1: Histograma del preu de cada període del mercat diari. L'eix horitzontal representa el preu (per MWh) en euros i el vertical el número de vegades que s'ha repetit aquest preu.

6.1 Sèries temporals

En aquest punt s'analitzaran els models obtinguts a partir de les sèries temporals. Aquests models han sigut els que finalment han predit tot l'any 2012 atès que s'adapten millor a les necessitats de l'estudi.

Com que per generar els models de la sessió 1 s'ha simplificat la metodologia, no es poden analitzar amb tant detall. Així doncs, si no es diu el contrari, els resultats que s'exposen corresponen als models del mercat diari.

Models

Seguint la metodologia que s'ha explicat en el punt 4.2, s'han creat els 12 models del mercat diari de l'any 2012. A l'hora d'identificar-los, s'ha observat com tots ells presenten un comportament molt semblant pel que fa a l'ACF i la PACF. És a dir, durant tot l'any les correlacions entre retards mantenen un patró molt constant. D'aquesta manera, la Figura 6.2 es repeteix, amb petites variacions, en tots els models estimats.

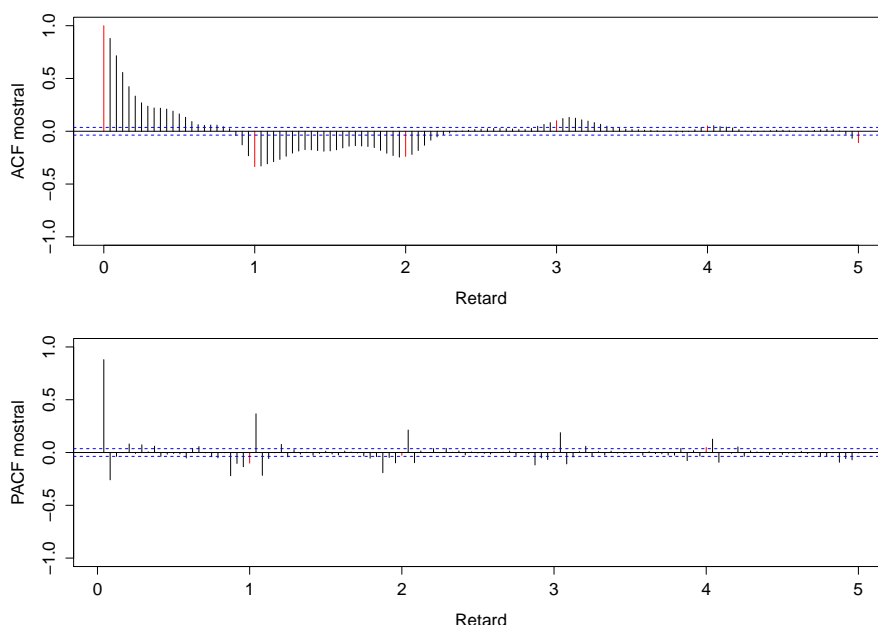


Figura 6.2: ACF i PACF mostrals utilitzats per identificar el model M1 del mercat diari.

D'aquesta manera, l'estructura identificada ha sigut la $SARIMAX(9,0,0)(0,1,2)_{24}$ i s'ha repetit en tots els models. A la Taula 6.2 s'han representat els models obtinguts. En aquesta s'ha indicat la variància del model, el període de generació, quantitat de valors atípics i els valors i desviacions dels coeficients (aquests últims entre parèntesi). També és important destacar que el temps d'estimació depèn, en gran mesura, de l'equip del qual es disposi. Tot i així, és verifica que en cap cas supera els 20 minuts.

Es pot veure com tots els estimadors de les variables significatives mantenen (en la majoria dels models) un ordre de magnitud semblant. S'entén per variable significativa aquella que el seu estimador sigui 2 cops més gran (en valor absolut) que el seu error estàndard, que s'ha expressat entre parèntesis.

La seva interpretació (tant per la part univariant com per les variables exògenes) és igual que la de un model lineal de regressió múltiple. Per tant, a partir del signe i la seva significació, es pot identificar si existeix una relació directa o inversa (amb la variable resposta). S'ha de tenir en compte que en la realització del codi s'han dividit les variables exògenes entre 10.000 atès que els coeficients resultats eren molt baixos. D'aquesta manera s'ha aconseguit

una millor resolució. A més, es recorda que a la variable resposta s'hi han aplicat logaritmes. Per tant, s'hauran de tenir present aquests aspectes en la seva interpretació quantitativa.

En el cas dels models del mercat diari, existeix una relació directe amb el preu d'una setmana enrere, la previsió de la demanda i el preu de tancament del mercat de futurs. En canvi, la previsió èlica manté una relació indirecta (mentre que la resta de variables no es consideren significatives).

Les variables exògenes dels models intradiaris mantenen les mateixes relacions amb la particularitat que la potència nuclear passa a ser significativa amb una relació inversa. A més, el mercat de futurs és més significatiu i no va perdent pes al llarg dels mesos.

Pel que fa a la part *AR* (regular) del mercat diari, s'observa que fins el retard 8, en general hi ha significació mentre que el número 9 deixa de ser-ho. En canvi, la part *MA* (estacional) sempre és molt significativa. Aquesta última té un gran pes sobre la predicció.

Atenent a les variables exògenes, és important tenir en compte que les magnituds no són comparables ni a la part *AR* ni a la *MA* ja que cadascuna va expressada amb unes unitats diferents. Per calcular la contribució de les variables als models, és calcularà el percentatge de la variabilitat que s'explica amb la seva incorporació.

En primer lloc es calcularà la contribució de les variables exògenes comparant la variació relativa de la variància del model univariant amb el segon model, que incorpora les variables exògenes (abans de linealitzar la sèrie). Per fer aquest càlcul s'utilitzen els models que defineixen *M2* atès que es considera que són un cas representatiu. Aquests es poden consultar en . El resultat d'aquest càlcul mostra com les variables exògenes expliquen un 5 % de la variància.

En canvi, si es fa aquest mateix càlcul comparant el model abans de linealitzar la sèrie amb el model que s'obté posteriorment, aquesta linealització explica pràcticament un 87 % d'aquesta. Per tant, es demostra com la sèrie del propi preu té un pes molt important. S'ha de tenir en compte que si es volgués comparar d'una manera estricta aquestes contribucions, s'hauria de fer partint del mateix model i amb una sola incorporació. En aquest cas es considera suficient atès que la diferència és evident.

Model	V. Atípics	Variancia	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6	ϕ_7	ϕ_8	ϕ_9	ϕ_{10}	Preu168	PDemanda	PEolica	Rhidro	Nuclear	OMP2D
M1	05/09/2011	03/01/2012	1,2794	-0,2611	-0,1087	-0,0369	-0,0165	0,0449	0,0248	-0,0131	0,0040	-0,8539	43,2301	0,2540	-0,3038	0,4226	-0,2767	14,3152
	129	0,0032	(0,0191)	(0,0305)	(0,0309)	(0,0310)	(0,031)	(0,0309)	(0,0309)	(0,0306)	(0,0188)	(0,0193)	(3,1956)	(0,0185)	(0,0352)	(0,4202)	(0,0943)	(7,8295)
	03/10/2011	31/01/2012	1,3357	-0,3594	-0,0803	-0,0505	0,0279	0,0559	-0,0141	-0,0244	0,0257	-0,8650	39,5927	0,2422	-0,3351	0,8244	-0,2768	14,8585
M2	111	0,0041	(0,0190)	(0,0313)	(0,0320)	(0,0321)	(0,0321)	(0,0320)	(0,0320)	(0,0313)	(0,0188)	(0,0185)	(3,3846)	(0,0197)	(0,0375)	(0,5058)	(0,1039)	(8,2228)
	01/11/2011	29/02/2012	1,2608	-0,2831	-0,0970	-0,0695	0,0618	-0,0265	0,0782	-0,0616	0,0292	-0,7967	37,4167	0,2472	-0,2900	0,3392	-0,0063	29,9439
M3	118	0,0050	(0,0190)	(0,0303)	(0,0307)	(0,0307)	(0,0308)	(0,0307)	(0,0307)	(0,0304)	(0,0188)	(0,0187)	(3,5512)	(0,0219)	(0,0352)	(0,6693)	(0,1121)	(9,1257)
	06/12/2011	04/04/2012	1,1065	-0,1476	-0,1560	-0,0151	0,0470	-0,0201	0,0642	-0,0068	-0,0018	-0,8175	29,4115	0,2628	-0,2329	0,3528	-0,1152	10,6596
	103	0,0030	(0,0190)	(0,0281)	(0,0281)	(0,0283)	(0,0282)	(0,0283)	(0,0281)	(0,0281)	(0,0188)	(0,0186)	(3,1992)	(0,0152)	(0,0237)	(0,5798)	(0,0902)	(8,7429)
M4	03/01/2012	02/05/2012	1,2192	-0,2058	-0,0890	-0,0809	-0,0435	0,0552	0,0843	0,0360	-0,0450	-0,5643	17,8920	0,3901	-0,2803	-1,7519	0,3136	23,6514
	110	0,0080	(0,0189)	(0,0297)	(0,0302)	(0,0302)	(0,0306)	(0,0305)	(0,0309)	(0,0319)	(0,0204)	(0,0214)	(4,3759)	(0,0281)	(0,0454)	(1,5544)	(0,2325)	(11,2713)
M5	31/01/2012	30/05/2012	1,1907	-0,1542	-0,1434	-0,0631	0,0558	-0,0607	0,0856	0,0481	-0,0188	-0,8293	19,9158	0,4212	-0,2996	0,3030	0,3628	30,0119
	117	0,0055	(0,0191)	(0,0292)	(0,0294)	(0,0294)	(0,0293)	(0,0294)	(0,0293)	(0,0292)	(0,0189)	(0,0195)	(3,9837)	(0,026)	(0,038)	(0,7039)	(0,1955)	(10,0502)
	06/03/2012	04/07/2012	1,1751	-0,1225	-0,0913	-0,1349	0,0699	-0,0888	0,1250	0,0630	-0,0445	-0,7729	14,7686	0,4959	-0,2549	0,1111	0,1572	33,7661
M6	93	0,0060	(0,0190)	(0,0290)	(0,0291)	(0,0290)	(0,0300)	(0,0302)	(0,0303)	(0,0302)	(0,0196)	(0,0197)	(4,4208)	(0,0296)	(0,0386)	(0,8208)	(0,2146)	(10,1312)
	03/04/2012	01/08/2012	1,2735	-0,2595	-0,0662	-0,1415	0,1295	-0,1449	0,1452	0,0405	-0,0233	-0,8674	21,6722	0,4597	-0,2832	0,6362	0,0739	38,3106
	86	0,0047	(0,0191)	(0,0305)	(0,0309)	(0,0307)	(0,0306)	(0,0306)	(0,0309)	(0,0306)	(0,0188)	(0,0194)	(4,1138)	(0,0294)	(0,0355)	(0,5516)	(0,1980)	(9,5222)
M7	01/05/2012	29/08/2012	1,0374	-0,1489	-0,0059	-0,0737	0,0183	-0,0359	0,0097	0,0987	-0,0143	-0,8327	28,2971	0,3307	-0,3024	0,4210	-0,2543	27,5211
	72	0,0020	(0,0191)	(0,0272)	(0,0274)	(0,0272)	(0,0272)	(0,0272)	(0,0272)	(0,0271)	(0,0190)	(0,0187)	(3,1841)	(0,0175)	(0,0202)	(0,2544)	(0,1006)	(7,0288)
M8	29/05/2012	26/09/2012	1,1793	-0,2189	-0,0467	-0,0824	-0,0155	-0,0383	0,1304	0,0512	-0,0544	-0,8454	29,3971	0,3049	-0,2775	-0,1450	-0,3751	15,2765
	104	0,0019	(0,0191)	(0,0292)	(0,0292)	(0,0292)	(0,0292)	(0,0292)	(0,0292)	(0,0290)	(0,0188)	(0,0185)	(3,3168)	(0,0178)	(0,0213)	(0,2110)	(0,1201)	(7,2366)
	03/07/2012	31/10/2012	1,3545	-0,3134	-0,1161	-0,0607	-0,0095	0,0397	0,0480	0,0927	-0,0908	-0,8751	30,0617	0,2974	-0,2114	0,2191	-0,4201	-7,1973
M9	89	0,0035	(0,0190)	(0,0316)	(0,0321)	(0,0322)	(0,0321)	(0,0321)	(0,0321)	(0,0315)	(0,0187)	(0,0190)	(3,7588)	(0,0232)	(0,0357)	(0,4249)	(0,3017)	(8,8640)
	31/07/2012	28/11/2012	1,3070	-0,2247	-0,1180	-0,0721	-0,0241	0,0428	0,0274	0,0862	-0,0725	-0,8487	25,5336	0,3509	-0,2599	0,6143	-0,3428	6,6376
	129	0,0043	(0,0191)	(0,0310)	(0,0313)	(0,0312)	(0,0313)	(0,0312)	(0,0311)	(0,0308)	(0,0187)	(0,0198)	(3,8195)	(0,0243)	(0,0397)	(0,6845)	(0,2557)	(7,7203)

Taula 6.2: Models obtinguts per al mercat diari.

Com s'ha comentat, tots els models utilitzats segueixen uns patrons semblants. De la mateixa manera, pràcticament tots ells tenen una variància i una quantitat de valors atípics del mateix ordre.

Tot i així s'ha de tenir en compte que no s'han utilitzat tots els models ja que s'ha seguit l'heurística explicada en el punt 4.2. A continuació es resumeix com ha evolucionat el model al llarg de l'any 2012 i quins models s'han utilitzat en cada període:

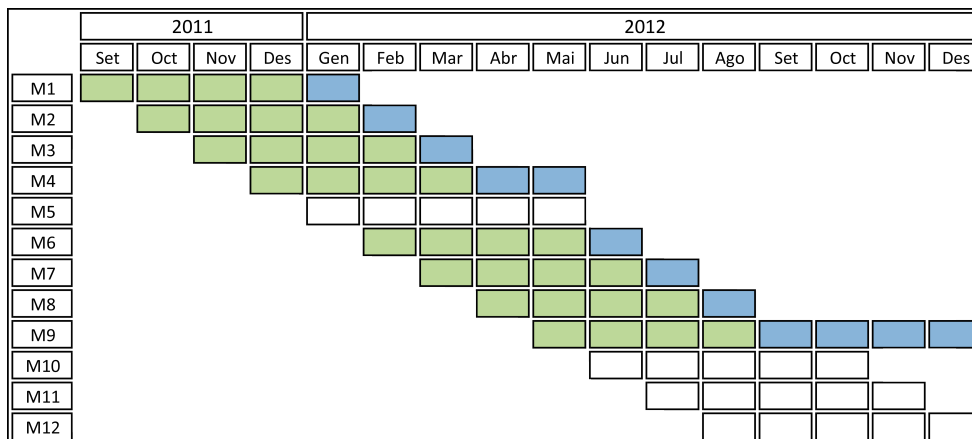


Figura 6.3: Resum de l'evolució dels models. De color verd el període de generació del model i de color blau el període que ha predit.

Pel que fa al mercat intradiari s'han trobat moltes similituds amb els que s'han exposat anteriorment. En primer lloc, l'estructura identificada ha sigut la $SARIMAX(9, 0, 0)(0, 1, 2)_{24}$. Tot i així, a la Taula 6.3 s'observa que possiblement, si s'hagués arribat fins el retard número 5 de la part AR regular, hagués sigut suficient.

En aquest cas, però, destaca l'elevat grau de significació del preu del mercat de futurs (OMIP) i que l'energia nuclear sembla que té una mica més de constància. La significació de la resta de variables té un comportament semblant als models del mercat diari.

Model	V. Atípics	Variancia	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6	ϕ_7	ϕ_8	ϕ_9	θ_{μ}	θ_{σ}	θ_{α}	Preu168	PDemanda	PEolica	Rhido	Nuclear	OMP2D
M1	05/09/2011	03/01/2012	1,2759	-0,2493	-0,0416	-0,1066	0,0536	-0,0489	0,0589	0,0017	-0,0059	-0,9048	(0,0191)	(0,0189)	41,5652	0,2980	-0,2714	1,4150	-0,2966	80,6207
	145	0,0048	(0,0192)	(0,0309)	(0,0310)	(0,0308)	(0,0309)	(0,0309)	(0,0309)	(0,0306)	(0,0189)	(0,0191)	(0,0189)	(0,0189)	(3,3769)	(0,0224)	(0,0470)	(0,5669)	(0,1202)	(9,7017)
M2	03/10/2011	31/01/2012	1,2843	-0,2576	-0,0664	-0,1212	0,0656	0,0284	0,0045	0,0123	-0,0140	-0,8688	(0,0187)	(0,0187)	35,5573	0,2860	-0,2937	1,5138	-0,2321	39,9093
	134	0,0057	(0,0190)	(0,0307)	(0,0310)	(0,0309)	(0,0310)	(0,0309)	(0,0309)	(0,0305)	(0,0188)	(0,0187)	(0,0187)	(0,0187)	(3,6009)	(0,0236)	(0,0490)	(0,7444)	(0,1269)	(9,9463)
M3	01/11/2011	29/02/2012	1,2362	-0,2424	-0,0271	-0,1124	0,0085	0,0480	0,0072	0,0284	-0,0236	-0,7604	(0,0184)	(0,0184)	35,8694	0,2856	-0,3274	2,1522	-0,0094	17,5528
	123	0,0086	(0,0190)	(0,0301)	(0,0303)	(0,0301)	(0,0303)	(0,0302)	(0,0302)	(0,0299)	(0,0188)	(0,0184)	(0,0184)	(0,0184)	(4,0324)	(0,0296)	(0,0533)	(1,1293)	(0,1502)	(12,7373)
M4	06/12/2011	04/04/2012	1,0506	-0,1509	-0,0667	-0,1059	0,0439	0,0645	-0,0130	-0,0226	0,0302	-0,8100	(0,018)	(0,018)	27,1090	0,2580	-0,2872	0,7372	-0,2478	-0,6590
	97	0,0059	(0,0189)	(0,0273)	(0,0274)	(0,0274)	(0,0275)	(0,0275)	(0,0276)	(0,0273)	(0,0189)	(0,0179)	(0,0179)	(0,0179)	(3,6542)	(0,0190)	(0,0278)	(0,6991)	(0,1200)	(11,041)
M6	31/01/2012	30/05/2012	1,0997	-0,1186	-0,0957	-0,0551	0,0247	0,0162	0,0000	0,0584	-0,0030	-0,8551	(0,0188)	(0,0188)	23,4504	0,4425	-0,3382	-0,1183	-0,1499	33,0027
	153	0,0066	(0,0190)	(0,0280)	(0,0281)	(0,0281)	(0,0281)	(0,0281)	(0,028)	(0,0279)	(0,0188)	(0,0188)	(0,0188)	(0,0189)	(3,6946)	(0,0275)	(0,0399)	(0,6701)	0,2130	(11,0523)
M7	06/03/2012	04/07/2012	1,1059	-0,1093	-0,1088	-0,0616	0,0274	0,0385	-0,0170	0,0518	0,0107	-0,7999	(0,0194)	(0,0194)	17,6308	0,4950	-0,3293	-0,9824	-0,3391	41,6036
	144	0,0063	(0,0191)	(0,0280)	(0,0282)	(0,0282)	(0,0290)	(0,0291)	(0,0291)	(0,029)	(0,0194)	(0,0197)	(0,0197)	(0,0197)	(3,9966)	(0,0293)	(0,0374)	(0,7601)	(0,2106)	(10,5390)
M8	03/04/2012	01/08/2012	1,2014	-0,2255	-0,0655	-0,0861	0,0289	0,0407	-0,0110	0,0567	0,0118	-0,8821	(0,0191)	(0,0191)	4,1877	0,5101	-0,3062	-0,5872	-0,4195	44,1896
	136	0,0045	(0,0190)	(0,0294)	(0,0298)	(0,0296)	(0,0297)	(0,0296)	(0,0296)	(0,0293)	(0,0189)	(0,0191)	(0,0191)	(0,0188)	(3,5755)	(0,0273)	(0,0336)	(0,5318)	(0,1924)	19,6802
M9	01/05/2012	29/08/2012	0,9102	-0,0608	-0,0281	-0,0547	0,0671	-0,0457	0,0232	0,0849	-0,0066	-0,8205	(0,0185)	(0,0185)	22,3773	0,3473	-0,2976	0,0198	-0,2954	19,6802
	86	0,0020	(0,0189)	(0,0254)	(0,0254)	(0,0254)	(0,0254)	(0,0254)	(0,0254)	(0,0254)	(0,0190)	(0,0187)	(0,0187)	(0,0187)	(3,0128)	(0,0166)	(0,0184)	(0,2393)	(0,1016)	(7,0151)

Taula 6.3: Models obtinguts per a la sessió 1 del mercat intradiari.

Prediccions

En primer lloc es comprova que el fet d'introduir el $MAPE_c$ no emmascari els resultats atès que no és la seva finalitat. Per fer-ho s'han comptabilitzat tots aquells que realment han modificat la mesura (és a dir, els que tenen un $MAPE$ diferent). Es recorda que aquesta magnitud expressa l'error comès per dia (és la mitjana horària de l'error relatiu expressada en tant per u).

D'aquesta manera s'han corregit 20 dies, que representen un 5,59 % dels dies predits. D'aquests, 10 són per la presència d'almenys una hora amb un preu de 0 €/MWh i que hagués implicat un error infinit. Es considera que aquest valor es baix i per tant, el $MAPE_c$ realitza correctament la seva funció. Per tant, la seva utilització està totalment justificada i alhora és un valor suficientment representatiu.

Per avaluar si aquesta metodologia és adequada, es considera interessant comparar la mitjana del $MAPE_c$ obtingut en cada període aplicant l'heurística enfront d'una estratègia on s'utilitzi cada model en el seu període previst. Aquesta comparativa s'ha representat a la Taula 6.4.

Amb la metodologia proposada la mitjana anual del $MAPE_c$ és de 12,99 % enfront del 14,42 % que s'obtidria inicialment.

Com s'observa a la Figura 6.3, hi ha 4 models que no s'han utilitzat ($M5$, $M10$, $M11$ i $M12$).

Període	Mitjana $MAPE_c$ amb heurística	Mitjana $MAPE_c$ sense heurística
1	0,0919	0,0919
2	0,1106	0,1106
3	0,1053	0,1053
4	0,1889	0,1889
5	0,1456	0,2887
6	0,1189	0,1189
7	0,1070	0,1070
8	0,1030	0,1030
9	0,1184	0,1184
10	0,1608	0,1253
11	0,1617	0,1740
12	0,1463	0,1984

Taula 6.4: Mitjana del $MAPE_c$ obtinguda per a cada període del mercat diari i per a cada estratègia.

Dels quatre models que no s'han utilitzat, només en un d'ells s'haurien obtingut millors prediccions. Es tracta del model $M10$, que hauria millorat la predicció en un 3,56 %.

Tot i així, el que realment destaca són els resultats obtinguts en el període 5. Inicialment s'hauria obtingut un $MAPE_c$ de 28,87 %, un 14,31 % més elevat que si s'aplica l'heurística.

Quan hi ha errors tan elevats, bàsicament hi ha dos possibles explicacions. Es tracta d'un període on hi ha una gran quantitat de valors atípics i/o el model ha sigut desvirtuat en la seva generació. Atès que amb el model $M4$ els errors són molt més baixos (pel mateix període analitzat), és molt probable que $M5$ s'hagi generat incorrectament degut a l'alt percentatge de valors atípics continguts en l'últim mes (període 4) de la seva generació. De fet, quan s'han analitzat els preus, s'ha vist que el període 4 té la mitjana més baixa de tot l'any i una quantitat considerable de valors atípics.

Es decideix avaluar diferents períodes amb diferents models per tal de clarificar aquesta situació.

Comparant el $MAPE_c$ de cada model (entre els períodes 1 i 4) s'observa que quan l'error és superior al 15 %, normalment ho és en tots (Figura 6.4). Aquest resultat sembla lògic atès que els models obtinguts són molt semblants.

També s'observa com al final del gràfic el $MAPE_c$ es dispara per tots els models. És important destacar que es tracta del període 4 i que a més, coincideix amb setmana santa on l'estacionalitat de la sèrie es pot veure totalment desvirtuada.

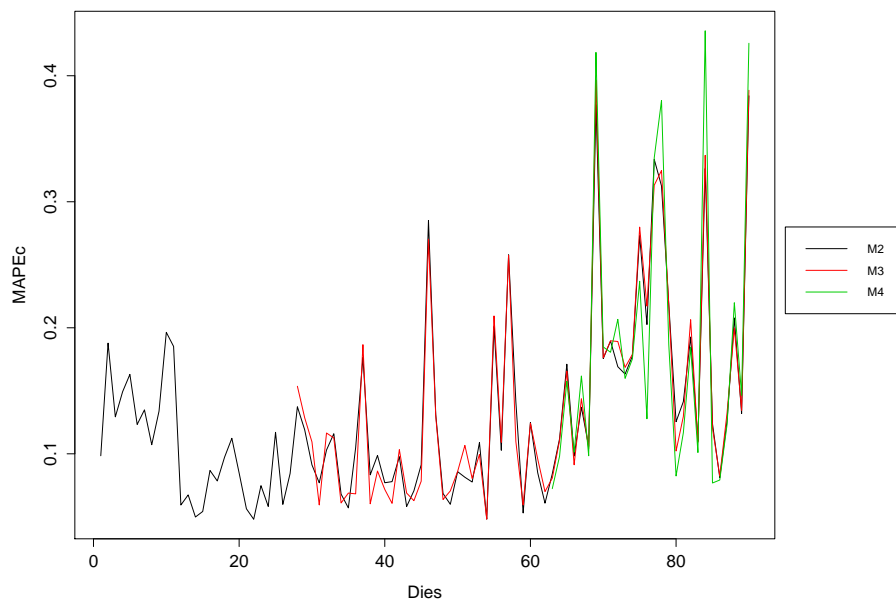


Figura 6.4: Representació del $MAPE_c$ obtingut per a cada model entre 01/02/2012 i 01/05/2012 (mercat diari).

A la Figura 6.5 s'ha representat novament el $MAPE_c$ de diferents models a partir del període 5. En aquest cas s'ha inclòs $M5$.

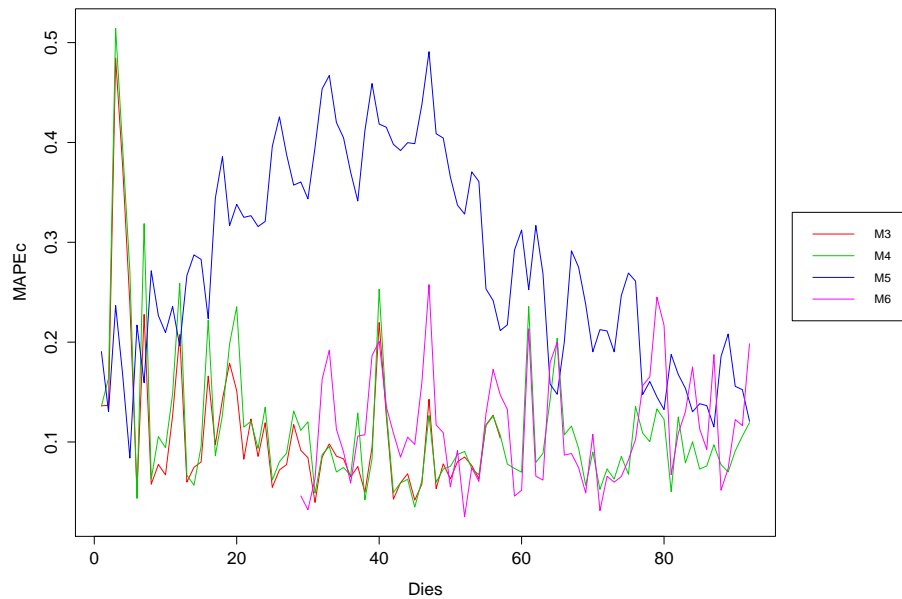


Figura 6.5: Representació del MAPEc obtingut per a cada model entre 02/05/2012 i 01/08/2012 (mercat diari).

Tots els models tenen un comportament semblant mentre que *M5* genera uns errors molt elevats (només en el tram final sembla que els seus errors es moderen). Per aquesta raó es descarta que el tram que s'ha predit contingui una quantitat excessiva de valors atípics.

S'observa que els coeficients que difereixen de la resta de models són de la part estacional. En tots els models utilitzats, la part *MA* d'ordre 24 és aproximadament de 0,8 mentre que en aquest cas (*M5*) és de 0,5. A més, la seva variància està per sobre de la resta (0,008).

Al final del període, l'error del model *M5* es va atenuant. És a dir, l'efecte estacional va perdent pes en la predicció. Una possible explicació és que es tracten dels mesos corresponents a l'estiu on els hàbits de consum es poden veure modificats.

Per últim es considera interessant analitzar els períodes amb uns errors més elevats (períodes 4, 5, 10, 11 i 12) que són els que tenen una mitjana del preu més baixa.

Si es representa l'energia diària per tecnologies en aquests períodes, s'observen unes puntes de producció en règim especial. Aquestes coincideixen amb períodes de preu 0 €/MWh i només es presenten en els períodes esmentats (a excepció del període 5).

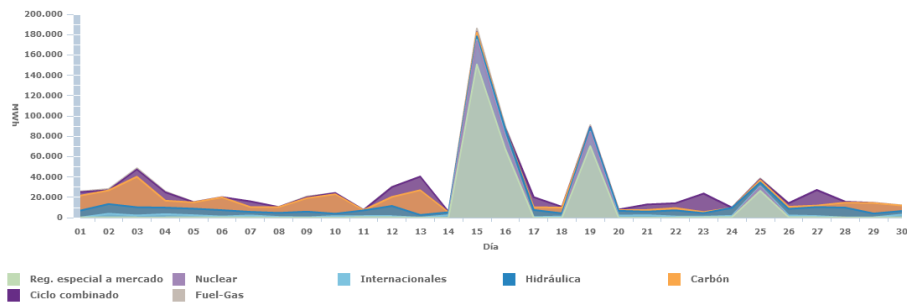


Figura 6.6: Producció diària d'energia per tecnologies del mes d'abril. Extret del lloc web d'OMIE.

Valors atípics i linealització de la sèrie

En el punt 6.1 s'ha vist com el fet de linealitzar la sèrie explica un 87 % de la variància. Tot i així, s'ha de tenir en compte que aquesta variància és del propi model.

Aquest fet podria induir a l'error atès que si es linealitza excessivament la sèrie, realment s'obtindria un model ideal que no s'ajustaria a la realitat. S'obtindria una variància molt baixa i un percentatge d'explicació elevat, però realment s'estaria empitjorant el model.

A més, tot i que es disposa d'un mecanisme de detecció i correcció dels valors atípics, aquests influeixen en la generació del model d'una manera determinant. La detecció i linealització es fa en base al model teòric del moment. Per tant, els valors atípics han influenciat en aquestes i en conseqüència, en la seva pròpia correcció.

A continuació es representen les sèries que han generat els models $M4$ i $M5$. Cal notar que hi ha un desfasament aproximat de 30 dies entre elles.

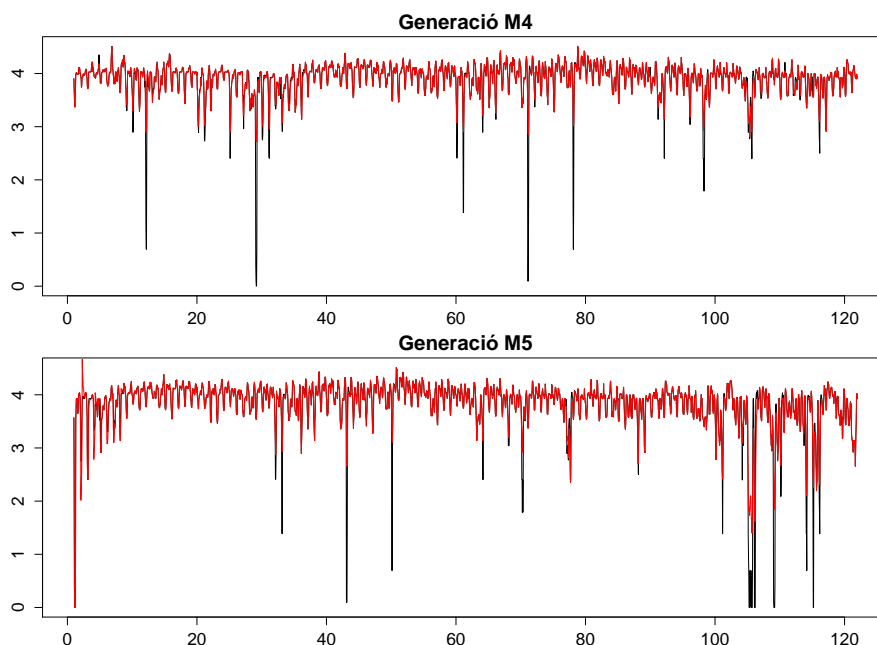


Figura 6.7: Sèries de preus de generació. L'eix horitzontal representen els dies i el vertical el logaritme del preu més una unitat. De color negre els valors considerats atípics.

Atenent als valors, els preus que generen $M5$ presenten molta volatilitat en l'últim mes (que coincideix amb el període 4). Aquest fet fa que alguns valors inicials no es considerin atípics.

En canvi, la sèrie que genera $M4$, presenta una major exigència. És a dir, considera alguns valors com atípics que $M5$ no considera.

Per tant, s'observa com linealitzar la sèrie no és una solució absoluta ja que els valors atípics influeixen en les correccions.

6.2 Xarxes neuronals

S'han predit els dos primers mesos de l'any 2012 (períodes 1 i 2) del mercat diari amb un model basat en xarxes neuronals. De la mateixa manera que amb els models de sèries temporals, s'han creat els equivalents a $M1$ i $M2$. S'ha fet així atès que són períodes amb un comportament poc atípic, podent comparar sense haver de fer cap suposició.

En aquest cas és important destacar que aquests models poden no ser òptims atès que existeixen mínims locals. Tot i així, en fer 20 repeticions es considera que el millor model aconseguir serà acceptable. Per tant, amb aquest anàlisi s'intentarà esbrinar si un model d'aquest tipus es podria adaptar bé però no permetrà extreure'n conclusions en un sentit absolut. A continuació s'adjunten els models resultants del primer període. En aquest cas el millor model és el que ha fet més iteracions, però no necessàriament ha de ser així:

	Error	Reached Threshold	Steps
3	95035.01	0.009462167	7750
12	96725.10	0.009105007	3112
1	99306.06	0.009318235	1132
6	111292.09	0.003559037	338
15	111909.09	0.009634984	360
9	113594.54	0.009739062	4430
16	116182.62	0.009759606	2251
2	121103.30	0.008452000	246
13	121675.93	0.009726857	2106
11	122426.04	0.007959658	168
17	133006.39	0.009777516	3526
4	133685.38	0.009417225	964
8	135647.84	0.009711549	4025
19	137375.86	0.002117085	170
10	138356.17	0.009409540	3568
14	138922.64	0.005636426	110
5	138922.64	0.003948149	83
18	138922.64	0.007955377	86
7	138922.64	0.005327167	92
20	138922.64	0.009477095	92

Figura 6.8: Models estimats amb les xarxes neuronals pel primer període.

A la Figura 6.9 s'ha representat el $MAPE_c$ obtingut amb les sèries temporals i amb les xarxes neuronals. S'observa com l'error de la predicció amb xarxes neuronals (12,91 %) és

superior a l'error amb sèries temporals (10,13 %). També destaca el gran error obtingut per les xarxes neuronals els primers dies però difícilment explicables per la topologia del model.

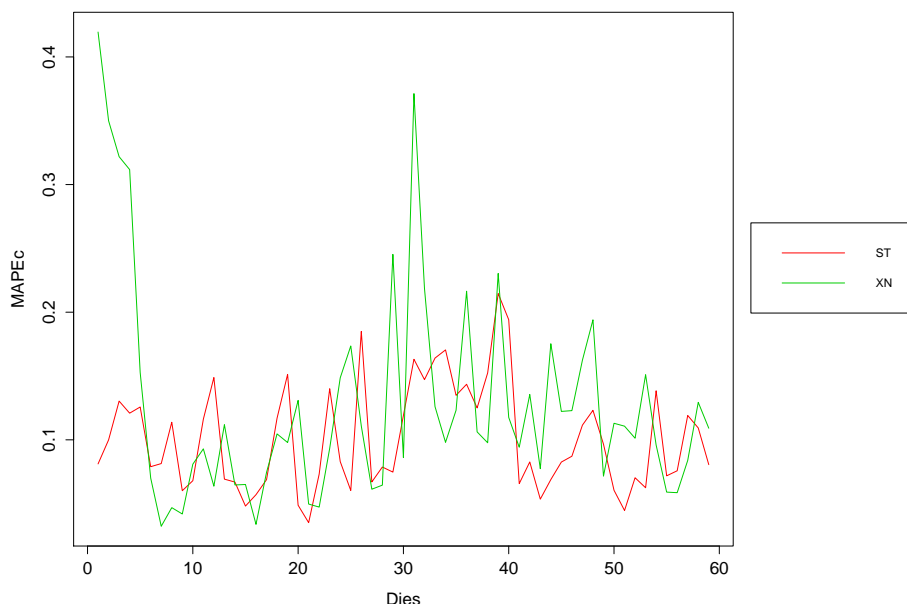


Figura 6.9: MAPEc per obtingut per cada model entre 04/01/2012 i 02/03/2012

És a dir, si es detecta algun comportament anòmal difícilment es pot explicar. Per aquesta raó l'estudi anual s'ha fet mitjançant les sèries temporals. A més, determinar la topologia del model no és una tasca trivial.

6.3 Selecció del mercat de compra

La metodologia que s'ha establert compara la predicció obtinguda pel mercat diari amb l'obtinguda per la primera sessió de l'intradiari. Aquestes prediccions es realitzen amb els models de sèries temporals. D'aquesta manera, es nomina l'energia per cada hora a la opció més favorable (el preu més barat).

Pel 2012 s'ha obtingut una matriu de 8592 caselles, on cadascuna indica si s'ha comprat en el mercat diari o en la sessió 1 de l'intradiari (a partir d'ara es denotarà per matriu de compra). Un any són 8760 hores però s'ha predit entre el 4/01/2012 i el 26/12/2012 per quadrar els dies de la setmana en la generació.

Atès que l'empresa sempre compra en el mercat diari, s'avalua la diferència entre aquesta opció i l'alternativa que es proposi. D'aquesta manera, per cada hora s'obté una nova matriu (matriu de diferències) resultat de restar-li a la matriu de preus reals del mercat diari, el preu (real) de la decisió que s'hagi pres. Així doncs, la casella podrà prendre 3 valors diferents:

- Valor positiu: S'ha decidit anar al mercat intradiari i el mercat diari a sigut més car. Per tant, opció favorable.
- Valor negatiu: S'ha decidit anar al mercat intradiari i el mercat diari a sigut més barat. Per tant, opció desfavorable.

- Valor nul: S'ha decidit mantenir-se en el mercat diari o bé ambdós mercats han marcat el mateix preu. Per tant, ni es guanya ni es perd.

S'han avaluat els resultats obtinguts durant l'any 2012, i s'observa que quasi el 50 % de les hores es compra en el mercat intradiari. D'aquestes, més de la meitat són una compra favorable (Taula 6.5). Tot i així, aquests resultats no es tradueixen en termes econòmics atès que no s'ha comprat la mateixa energia en totes les hores. De fet, en volum d'energia, les compres favorables representen un 25,10 % del total mentre que les desfavorables representen un 14,70 %.

Tipus de compra	Percentatge
Favorable	26,85%
Desfavorable	15,71%
Nul·la	6,68%
No compra	50,76%

Taula 6.5: Resultats obtinguts aplicant la metodologia en funció de les hores de participació.

Per tenir en compte aquesta implicació econòmica, es ponderarà l'estalvi de cada hora segons la corba de càrrega tipus A del 2012 (adaptada a les 8592 hores). Aquest perfil correspon a tarifes amb un sol període (és a dir, les 2.0A i les 2.1A) i representa aproximadament un 80 % dels clients d'HolaLuz.com. És a dir, es e en compte quan energia s'ha comprat en cada franja horària.

Per tant, la suma de tots aquests estalvis ponderats serà l'estalvi per cada MWh que es compri seguint aquesta metodologia.

Sumant els termes positius s'obté que amb les compres favorables s'estalvia 0,8210 €/MWh mentre que sumant les compres desfavorables s'obté un sobrecost de 0,3980 €/MWh. Per tant, s'observa com globalment s'obté un estalvi de 0,4230 €/MWh participant, en volum d'energia, un 45,76 %.

Per tant, utilitzant aquesta metodologia s'estaria estalviant pràcticament un 17 % del marge brut.

Es considera interessant calcular quin hauria sigut l'estalvi o el sobrecost de comprar tota l'energia en el mercat intradiari. Val a dir que es tracta d'un cas pràcticament impossible ja que l'empresa no es plantejaria en cap cas una estratègia d'aquest tipus. En aquest cas s'hauria fet una compra favorable un 52,74 % dels cops i una desfavorable un 33,38 %. Per tant, tot i que no te per què tenir una implicació econòmica directe, la metodologia millora el rati d'encerts.

Fent la nova matriu de diferències i ponderant-la per la corba de càrrega s'obté un estalvi anual de 0,7317 €/MWh. Per tant, s'estaria estalviant més que amb la metodologia però participant amb un 100 % del volum.

Tot i així, s'observa que el rati d'estalvi en funció del volum gestionat és millor en el cas de la metodologia.

Capítol 7

Estudi econòmic

L'objectiu principal d'aquest projecte es crear una eina que permeti escollir en quin mercat s'ha de nominar l'energia per minimitzar-ne el cost d'adquisició. Per tant, per fer una valoració global d'aquest, és necessari realitzar l'estudi econòmic. Posteriorment, amb els resultats que s'obtinguin es podrà determinar la viabilitat del projecte.

Per a fer aquest estudi, en primer lloc s'avaluaran tots els costos associats al projecte i posteriorment s'analitzarà la viabilitat d'aquesta inversió. També es tindran en compte possibles costos de manteniment de l'eina i es contemplaran diferents escenaris de reducció de costos.

7.1 Desenvolupament de l'eina

Tenint en compte que s'utilitza un software lliure, l'únic cost associat a la realització del projecte són les hores dedicades per part de l'enginyer i el hardware utilitzat.

L'enginyer (analista), realitza una dedicació de 540 hores amb una remuneració de 40 €/h. Aquesta inclou IRPF, seguretat social i quotes patronals. Per tant, el cost total és de 21.600 €.

Per realitzar aquest estudi, s'utilitza un ordinador valorat en 1.000 €. Es considera que la vida útil d'aquest és de 5 anys i s'utilitza durant 6 mesos. Per tant, el seu cost associat és de 100 €.

Sumant aquests dos valors s'obté una inversió inicial de 21.700 €. Posteriorment es tenen en compte despeses generals, costos indirectes i costos imprevistos.

Es resumeix el cost de la inversió en la Taula 7.1.

Concepte	Unitat	Preu unitari	Quantitat	Total
Enginyer	hores	40 €	540	21.600 €
Hardware informàtic	unitats	1.000 €	0,1	100 €
				21.700 €
Despeses generals	15%			3.255 €
Costos indirectes	10%			2.170 €
Costos imprevistos	10%			2.170 €
			TOTAL	29.295 €

Taula 7.1: Cost de la inversió.

7.2 Anàlisi de la inversió

Per analitzar la inversió es prendrà un horitzó de 5 anys. Com que s'haurà de generar el codi, es considera que l'any 2012 no hi hauria cap estalvi (tampoc cap cost que no sigui la pròpia inversió). Per tant, hagués entrat en funcionament a partir de l'any 2013.

Tot i que un cop generat el codi, l'eina és totalment automàtica, aquesta requerirà d'un tècnic que la supervisi. La remuneració d'aquest és de 20 €/h i realitzarà una dedicació de 15 hores anuals. Per tant, el cost de manteniment per any serà de 300 €. Aquest inclou IRPF i seguretat social del tècnic.

Per fer aquest anàlisi, es tindran en compte les previsions de creixement que té l'empresa i s'avaluaran diferents escenaris.

El primer cas avaluarà la situació actual de l'empresa. Actualment es compra tota l'energia al mercat diari de manera que l'estalvi és directament 0,4230 €/MWh. Aquest estalvi representa aproximadament un 17 % del marge brut.

El segon cas considerarà que l'empresa compra un 80 % de l'energia al mercat diari i el 20 % al mercat intradiari. Es considera aquest un escenari que l'empresa podria prendre de manera "intuitiva". Aquesta fórmula generaria un estalvi respecte comprar sempre en el mercat diari de 0,1463 €/MWh. Per tant, l'estalvi que aportarà l'eina serà de 0,2766 €/MWh. Aquest representa un 11 % del marge brut.

Per últim, es considerarà que l'empresa utilitza una estratègia agressiva, comprant el 50 % de l'energia en el mercat intradiari i la resta al diari. Aquesta estratègia generaria un estalvi respecte comprar sempre en el mercat diari de 0,3658 €/MWh. Per tant, l'eina aportaria un estalvi de 0,0571 €/MWh. Aquest estalvi representa aproximadament un 2 % del marge brut. A més, és important tenir en compte que l'eina participa en el mercat intradiari un 45 % en volum. Per tant, participa amb menys volum i aporta més estalvi.

Participar al mercat intradiari amb un volum raonable d'energia es considera interessant, donat que la liquiditat d'aquest mercat és menor que la del diari i, per tant, el risc i la volatilitat de l'intradiari són més altes.

En la Taula 7.2 es mostra l'estalvi generat per cada cas i per a cada any, tenint en compte les previsions de vendes de l'empresa.

Any	Energia venuda	Estalvi cas 1	Estalvi cas 2	Estalvi cas 3
2012	60 GWh	25.379 €	16.599 €	3.429 €
2013	100 GWh	42.298 €	27.664 €	5.715 €
2014	180 GWh	76.136 €	49.796 €	10.287 €
2015	300 GWh	126.893 €	82.993 €	17.145 €
2016	450 GWh	190.339 €	124.490 €	25.717 €
2017	600 GWh	253.785 €	165.987 €	34.290 €

Taula 7.2: Estalvi per a cada cas i any.

Com que la inversió del projecte s'ha finançat amb fons propis de l'empresa, no es considera cap tipus d'interès. Així doncs, ja es disposa de la informació necessària per avaluar aquesta inversió per a cada cas.

Cas 1

Pel primer cas, els resultats de la Taula 7.3 mostren com es recuperaria la inversió abans de finalitzar el 2013 i amb una TIR (taxa de retorn intern) de 206 %. Tenint en compte que una inversió es considera viable si la TIR és superior al 10 %, és evident que aquesta inversió és totalment viable.

	Any 0	Any 1	Any 2	Any 3	Any 4	Any 5
	2012	2013	2014	2015	2016	2017
Inversió inicial	29.295 €					
Estalvi anual eina		42.298 €	76.136 €	126.893 €	190.339 €	253.785 €
Costos operatius		300 €	300 €	300 €	300 €	300 €
Flux de caixa	-29.295 €	41.998 €	75.836 €	126.593 €	190.039 €	253.485 €
Total acumulat	-29.295 €	12.703 €	88.538 €	215.131 €	405.169 €	658.655 €
Taxa interna de retorn	206%					

Taula 7.3: Anàlisi de la inversió per al primer cas.

Cas 2

Pel que fa al segon cas (Taula 7.4), la inversió es recupera a principis del 2014 i s'obté una TIR de 149 %. Per tant, en aquest hipotètic cas també seria viable realitzar la inversió.

	Any 0	Any 1	Any 2	Any 3	Any 4	Any 5
	2012	2013	2014	2015	2016	2017
Inversió inicial	29.295 €					
Estalvi anual eina		27.664 €	49.796 €	82.993 €	124.490 €	165.987 €
Costos operatius		300 €	300 €	300 €	300 €	300 €
Flux de caixa	-29.295 €	27.364 €	49.496 €	82.693 €	124.190 €	165.687 €
Total acumulat	-29.295 €	-1.931 €	47.566 €	130.259 €	254.449 €	420.136 €
Taxa interna de retorn	149%					

Taula 7.4: Anàlisi de la inversió per al segon cas.

Cas 3

Per últim, s'analitza el tercer cas, que és el més desfavorable. La Taula 7.5 mostra com la inversió es recuperaria l'any 2015 i s'obté una TIR de 38 %. Per tant, la inversió continua essent totalment viable.

	Any 0	Any 1	Any 2	Any 3	Any 4	Any 5
	2012	2013	2014	2015	2016	2017
Inversió inicial	29.295 €					
Estalvi anual eina		5.715 €	10.287 €	17.145 €	25.717 €	34.290 €
Costos operatius		300 €	300 €	300 €	300 €	300 €
Flux de caixa	-29.295 €	5.415 €	9.987 €	16.845 €	25.417 €	33.990 €
Total acumulat	-29.295 €	-23.880 €	-13.893 €	2.952 €	28.369 €	62.358 €
Taxa interna de retorn	38%					

Taula 7.5: Anàlisi de la inversió per al tercer cas.

Per tant, amb els resultats obtinguts es confirma la viabilitat econòmica del projecte en qualsevol dels casos exposats. Si es té en compte que la situació actual de l'empresa és el primer cas, és evident que es tracta d'una inversió molt recomanable que implicaria, a finals del 2017, un flux acumulat de 658.655 €.

Capítol 8

Impacte ambiental i social

En aquest projecte s'ha dissenyat una eina que permet escollir en quin mercat s'ha de nominar l'energia per minimitzar-ne el cost d'adquisició. Per tant, el seu funcionament es basa en el diferencial de preu que hi ha entre mercats. D'aquesta manera, quan l'eina indica que el mercat intradiari serà més barat, es comprarà en aquest. Altrament, es comprarà l'energia en el mercat diari, tal i com ha fet l'empresa fins ara.

Si es compara amb la situació inicial, cada cop que HolaLuz.com participa en la primera sessió del mercat intradiari, està modificant la corba de la demanda agregada d'ambdós mercats. En el cas del mercat diari la desplaça cap a l'esquerra mentre que en el mercat intradiari la desplaça en sentit contrari. Aquest fet és important atès que cada vegada que l'eina encerti, farà que la diferència de preu entre aquests mercats es redueixi (sempre tenint en consideració que el volum del mercat que posseeix HolaLuz.com és relativament baix).

Aquesta reducció de la diferència de preus és important atès que un preu elevat implica, en la major part dels casos, l'entrada en funcionament de centrals de carbó i cicles combinats de baixa eficiència. Aquestes entren en el mercat a un preu elevat atès que el MWh produït és més car. Per tant, tota acció que redueixi la demanda d'un mercat diari amb un preu elevat, estarà reduint la participació d'aquest tipus de centrals.

Aquest fet és cert sempre i quan la predicció de l'eina sigui correcta. En aquest estudi s'ha vist com les compres favorables representen, en volum d'energia, un 25,10 % del total mentre que les desfavorables representen un 14,70 %. Per tant, aquesta eina està ajudant a reduir la participació de centrals poc eficients i que en molts casos, utilitzen combustibles que emeten diòxid de carboni en la seva combustió.

Per contra, en el mercat intradiari s'estarà incrementant el preu. Tot i així, en tractar-se d'un preu més baix, no necessàriament han d'entrar en funcionament centrals menys eficients ja que treballen per franges de preu elevades. Per aquesta raó, el fet de reduir la diferència de preus entre els mercats és un aspecte positiu.

Atenent a l'aspecte econòmic, l'empresa estarà comprant l'energia a un preu més baix. Aquest fet permetrà oferir uns preus finals més competitius i s'estarà fomentant la competència en el sector. Per tant, repercutirà directament sobre la societat.

Alhora, aquest estalvi permetrà incrementar el marge comercial, fet que es tradueix en una major seguretat financera. Aquest aspecte és clau per tota empresa jove, i més en aquest sector. S'ha de tenir present que es tracta d'un mercat extremadament volàtil. Per aquesta raó, en els contractes (amb els clients finals) on s'han acordat preus fixes, existeix la possibilitat que HolaLuz.com vengui, en alguns períodes, energia per sota del preu de cost.

Aquest fet no suposa un problema per les grans empreses que dominen el sector, però per les petites comercialitzadores pot suposar la fallida. La conseqüència final és una falta de competència en el sector. Així doncs, no només es fomentarà la competència del sector sinó s'estarà garantint que aquesta competència tingui una continuïtat.

En resum, la reducció del cost d'adquisició no només afavorirà a la pròpia empresa sinó que repercutirà sobre la societat. Un ventall més ampli de possibilitats on escollir, una major competència i sobretot, una garantia de supervivència d'aquesta competència.

Tot i que no és una conseqüència directe d'aquest projecte, és important destacar que HolaLuz.com només comercialitza energia renovable certificada per la CNMC (*Comisión nacional de los mercados y la competencia*). Per poder comercialitzar aquesta energia, l'empresa ha de comprar aquests certificats directament als productors. Per tant, amb el creixement d'HolaLuz.com, s'està incentivant d'una manera molt directa l'ús d'energies renovables. A més, l'empresa ha creat, des de la seva fundació, 45 llocs de treball qualificats, un 90 % dels quals tenen titulació superior i amb una mitjana d'edat per sota dels 30 anys.

Conclusions i futures línies de treball

L'objectiu d'aquest projecte és reduir el cost d'adquisició de l'energia elèctrica per part d'HolaLuz.com escollint, en cada moment, el mercat més favorable. Aquest objectiu s'assoleix mitjançant una eina que compara les prediccions puntuals a curt termini de cada mercat i escull aquell amb una predicció de preu menor. Dels tres possibles escenaris que s'han plantejat, el més favorable (i que és l'actual) aconsegueix un estalvi de 0,4230 €/MWh, que representa (aproximadament) un 17 % del marge comercial mentre que en el més desfavorable és un estalvi d'un 2 %. Després de fer l'estudi econòmic d'aquest s'observa que la inversió és totalment viable.

Tot i que no té una implicació econòmica directa, des del punt de vista estadístic s'ha vist com es realitzen més compres favorables (aproximadament un 26 %) que desfavorables (aproximadament un 15 %). Per tant, aquesta estratègia millora el rati d'encerts respecte comprar sempre en la primera sessió del mercat intradiari.

Els resultats mostren com els models de sèries temporals s'adapten millor a les necessitats del projecte per la seva simplicitat, baix cost computacional i unes prediccions acceptables. En canvi, els models de xarxes neuronals han tingut uns errors lleugerament superiors. Tot i així, no s'han descartat aquests últims per les prediccions sinó per la dificultat per detectar i adaptar errors i analitzar la situació canviant del mercat.

Tant pel mercat diari com per l'intradiari s'ha determinat que el millor model correspon a un $SARIMAX(9, 0, 0)(0, 1, 2)_{24}$ en tots els períodes. Per tant, la correlació de la sèrie de preus manté un patró constant durant tot l'any. De fet, els coeficients de la part *AR* regular i *MA* estacional han mantingut un ordre de magnitud molt constant en tots els models (a excepció de *M5*, on s'ha detectat ràpidament el problema).

Les prediccions que s'han fet amb aquests models basats en sèries temporals han tingut un error anual ($MAPE_c$) d'un 12,99 % en el cas del mercat diari i d'un 13,14 % en el cas del mercat intradiari. Aquest error es considera raonable atès que es tracta d'un mercat amb una elevada volatilitat. A més, l'heurística utilitzada per validar els models del mercat diari, ha evitat utilitzar models molt desvirtuats degut a la presència de valors atípics. A partir dels resultats observats, es fa evident que si s'hagués escollit un criteri on es valorés el $MAPE_c$ d'un instant posterior al de generació i aquest coincidís amb un valor atípic, es podria acceptar un model totalment desvirtuat. Per aquesta raó, tot i que el temps d'estimació d'aquests models en cap cas supera els 20 minuts (ajustant-se als objectius especificats), l'actualització no és intensiva.

Per aquesta raó, tot i que el temps d'estimació en cap cas supera els 20 minuts, no s'actualitzen els models de forma intensiva

Pel que fa al temps d'estimació d'aquests models, en cap cas es supera els 20 minuts de

manera que s'ajusten als objectius especificats.

Pel que fa a les variables exògenes, s'observa com les que tenen un pes important en el model (preu d'una setmana enrere, previsió de la demanda i previsió de l'energia eòlica) són aquelles amb valors horaris. Val a dir que aquestes dades també són les més evidents i fiables atès que provenen de REE o de la sèrie del propi preu. D'aquestes tres, l'única millora que es proposa és linealitzar el preu d'una setmana enrere ja que en estar com a variable exògena no s'ha corregit. És possible que aquest fet ajudi a millorar el model.

Tot i així, s'observa com la major part del pes en el model es concentra en la part univariant. Aquesta explica un 87 % de la variància del model enfront del 5 % que expliquen les variables exògenes. Per aquesta raó resulta clau la detecció i correcció dels valors atípics ja que afecten directament a la part univariant del model. Tot i així, s'ha de vigilar atès que una correcció excessiva de la sèrie tindria un efecte contrari al desitjat.

Atès que una part molt important dels models es basa en la part univariant, aquests són incapaços de detectar canvis bruscos. Per aquesta raó es comporten d'una manera molt adequada en períodes estables però s'obtenen unes prediccions molt errònies per canvis puntuals del preu. Així doncs, són uns models que pateixen per valors baixos del preu.

Per tal de millorar el model es proposen algunes alternatives com utilitzar la variable potència nuclear per diferències, actualitzar les dades hidrològiques diàriament amb l'ajuda de l'aplicació SAIH (*Sistema Automático de Información Hidrológica*) de MAGRAMA o crear un històric propi del valor que pren el mercat de futurs (ja que actualment s'utilitza el preu de tancament de dos dies abans).

Futures línies de treball

Es considera que els models tenen marge de millora, raó per la qual s'han esmentat algunes alternatives o propostes. A continuació es detallen aquests possibles estudis:

- Complementar l'estudi amb un model GARCH. Fins ara s'han fet estimacions puntuals del preu, però es podria complementar amb un interval de confiança que tingui en compte els períodes d'alta volatilitat.
- Realitzar el mateix estudi per la resta de mercats intradiaris (incorporant noves variables si és necessari). D'aquesta manera es podria plantejar comprar l'energia en diferents mercats intradiaris.
- Obtenir dades diàries de les reserves hidroelèctriques. Es proposa analitzar el *Sistema Automático de Información Hidrológica* de MAGRAMA per tal de disposar de dades a nivell diari. També es considera important analitzar les centrals reversibles.
- Analitzar noves possibles variables que no s'hagin inclòs en els models.
- Mesurar la qualitat dels valors atípics i aprofundir en la seva interpretació. En el criteri d'actualització tots els atípics han tingut el mateix pes. Es considera interessant analitzar el grau d'influència per millorar la metodologia.
- Creació d'històrics propis. En cas del preu del mercat de futurs només s'ha disposat del valor de tancament del dia abans del mercat diari. Aquesta variable presenta un comportament peculiar i es considera interessant disposar d'informació més detallada del seu comportament
- Realitzar l'estudi amb les xarxes neuronals adequant les variables a les seves característiques.

Bibliografia

- [1] DEREK W.B. and KARAKATSANI N. Forecasting electricity prices. 2003. 21, 27
- [2] GIRISH G.P. Modeling and forecasting day-ahead hourly electricity prices: A review. *Proceedings of the International Conference on Business Management and Information Systems*, 2012. 21
- [3] AGGARWAL S.K., SAINI L.M., and KUMAR A. Electricity price forecasting in deregulated markets: A review and evaluation, 13-22. *International Journal of Electrical Power and Energy Systems*, 31, 2009. 25
- [4] GONZÁLEZ A.M., ROQUE A.M.S., and GARCÍA-GONZÁLEZ J. Modeling and forecasting electricity prices with input/output hidden markov models. *IEEE*, 2005. 27
- [5] BOSSY M., MAÏZI N., OLSDER G.J., POURTALLIER O., and TANRÉ E. *Dynamic games: theory and applications. Electricity prices in a game theory context 135-159*. Springer, 2005. 28
- [6] WERON R. *Modeling and forecasting electricity loads and prices: A statistical approach*. John Wiley and Sons, 2007. 28
- [7] ANZBAZHAGAN S. and KUMARAPPAN N. Day-ahead deregulated electricity market price forecasting using recurrent neural network. *IEEE*, 2012. 29
- [8] PEÑA SÁCHEZ D. *Modelos y Métodos. Tomo II: Modelos Lineales y Series Temporales*. Alianza Universidad Textos, 1987. 31
- [9] MARTÍ RECOBER M. and MUÑOZ GRÀCIA M.P. *Previsió i sèries temporals. Mètodes empírics, models ARIMA, metodologia i casos*. Departament d'Estadística i Investigació Operativa, UPC, 2001. 31
- [10] BOX G.E.P., JENKINS G.M., and REINSEL GREGORY C. *Time series analysis: forecasting and control*. Prentice-Hall International Inc, 3rd edition, 1994. 35, 36, 38
- [11] RED ELÉCTRICA DE ESPAÑA. *e-sios*. <http://www.esios.ree.es>, 2 de desembre del 2013. 39
- [12] CONSEJO DE SEGURIDAD NUCLEAR. *Estados operativos*. <http://www.csn.es>, 5 de desembre del 2013. 40
- [13] CASADO DE AMEZÚA A.O. Anales de mecánica y electricidad. análisis de la gestión hidroeléctrica en el sistema español 20-27. *Asociacion de Ingenieros del ICAI*, 2003. 40
- [14] ALIMENTACIÓN Y MEDIO AMBIENTE MINISTERIO DE AGRICULTURA. *Boletín Hidrológico*. <http://portal.magrama.gob.es/BoleHWeb>, 16 de desembre del 2013. 40

- [15] OPERADOR MERCADO IBERICO POLO PORTUGUÉS. *Bolsa de derivados del MIBEL*. <http://www.omip.pt>, 28 de novembre del 2013. 40
- [16] OPERADOR DEL MERCADO IBERICO ESPAÑOL. <http://www.omie.es> 25 de novembre del 2013. 40
- [17] SHUMWAY R.H. and STOFFER D.S. *Time Series Analysis and Its Applications with R examples*. Springer, 2011. 41
- [18] MARTIN DEL BRÍO B. and SANZ A. *Redes Neuronales y Sistemas Borrosos*. RA-MA, 2001. 48

Annexos

Annex A

Sèries temporals: Cas pràctic

Aquest annex explica, pas a pas, com s'obté un model pel mercat diari. Es recorda que en el cas del mercat intradiari es simplifiquen alguns passos.

A.1 Determinació del model univariant

Introducció de les dades i transformacions prèvies

En primer lloc s'introdueix la sèrie temporal de preus. Com s'ha comentat, es tracta d'una finestra de 121 dies on la sèrie rep el nom de l'últim mes inclòs. Cada dia està format per 24 observacions que representen les 24 hores del dia. En aquest cas, la sèrie inclou l'últim dia de gener de l'any 2012 (31 de gener) i els 120 dies anteriors. Per tant, és la sèrie que generarà el model *M2*.

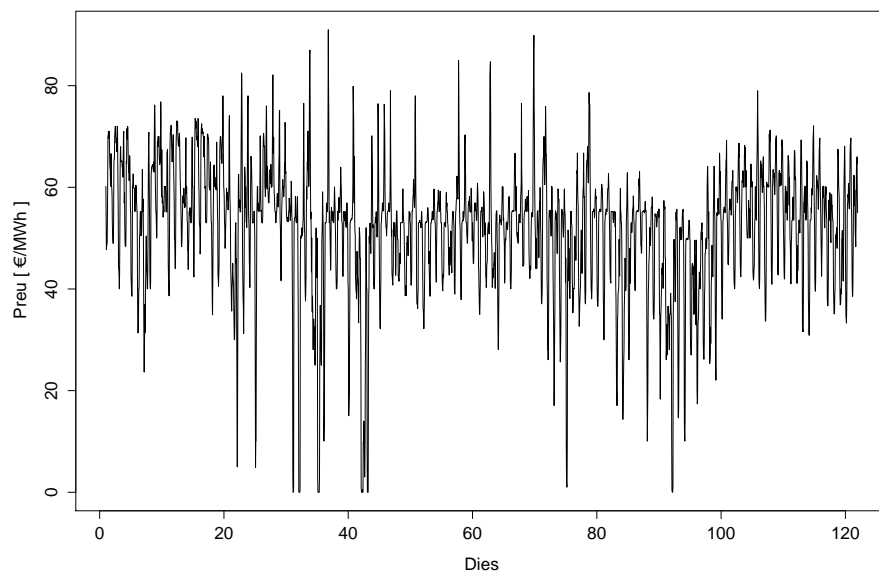


Figura A.1: Representació gràfica dels preus que generen *M2*.

Tal i com s'ha comentat en el procediment, aquesta sèrie es transforma mitjançant logaritmes. Així doncs, es fa el logaritme de la sèrie original incrementada en una unitat (Figura A.2). S'observa com les observacions s'han allunyat (relativament) de la cota inferior. Aquest fet permet detectar fàcilment els atípics i evitar canvis relatius molt elevats.

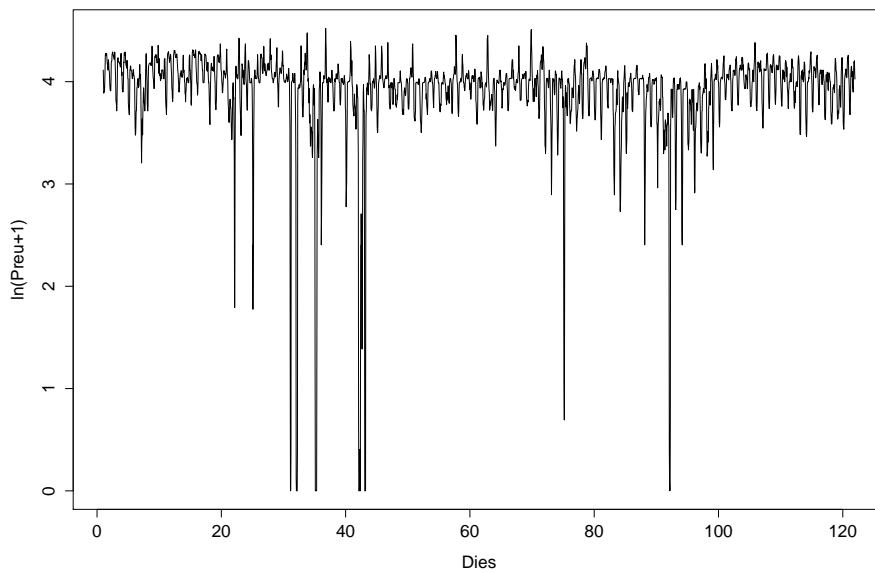


Figura A.2: Representació gràfica del logaritme neperià de la sèrie.

Tot i que presenta una alta volatilitat, s'observa un patró de comportament (Figura A.3). Aquest es repeteix cada 24 observacions i representa l'estacionalitat diària ($s = 24$).

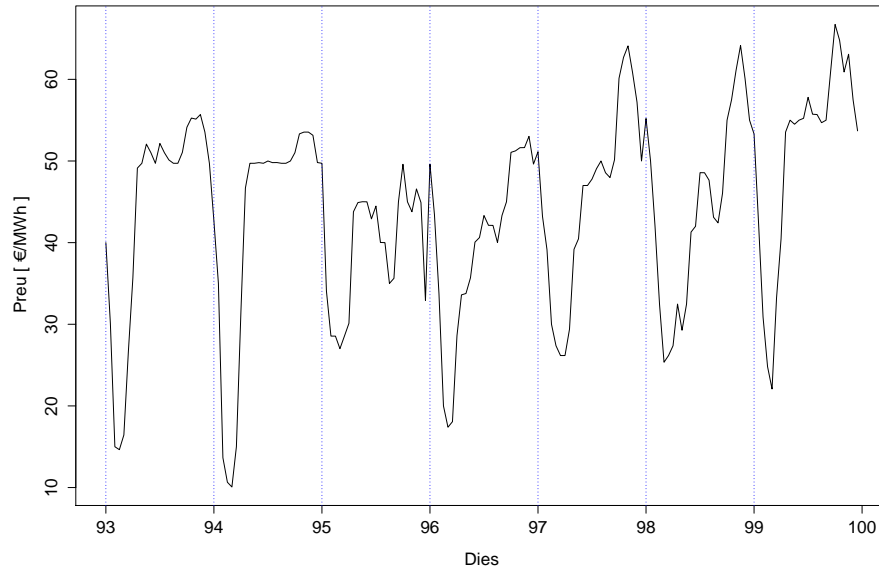


Figura A.3: Representació de la setmana 14 de la sèrie que genera M2

Per aquesta raó s'aplica una diferència d'ordre 24, corregint l'efecte estacional de la sèrie i estabilitzant la mitjana (Figura A.4). S'ha de tenir en compte que quan s'apliquen diferències es perden $d + sD$ observacions. Així doncs, de la sèrie de $120 \cdot 24$ observacions es perdran les 24 primeres ($d = 0$, $s = 24$ i $D = 1$).

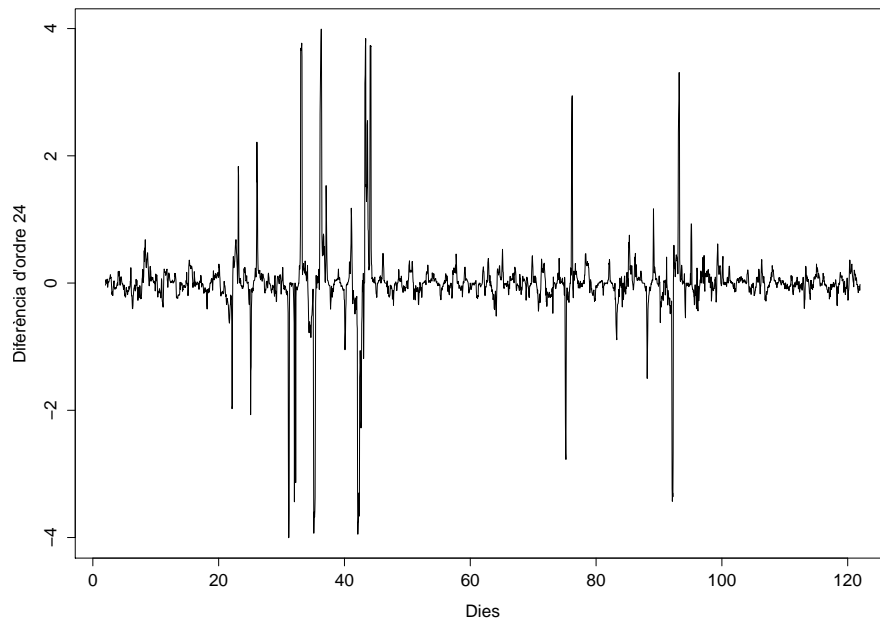


Figura A.4: Sèrie original amb una diferència d'ordre 24.

Com s'ha comentat, si s'apliquen diferències després d'aplicar logaritmes, s'està treballant

amb increments relatius (rendiments). En el cas d'estudi s'estan utilitzant rendiments diaris.

Anàlisi de l'ACF i la PACF mostrals

Un cop aplicades les diferències, la sèrie ja té les propietats necessàries per estimar un model $SARIMA(p, d, q)(P, D, Q)_s$. Per fer-ho, s'utilitza l'ACF i la PACF mostrals:

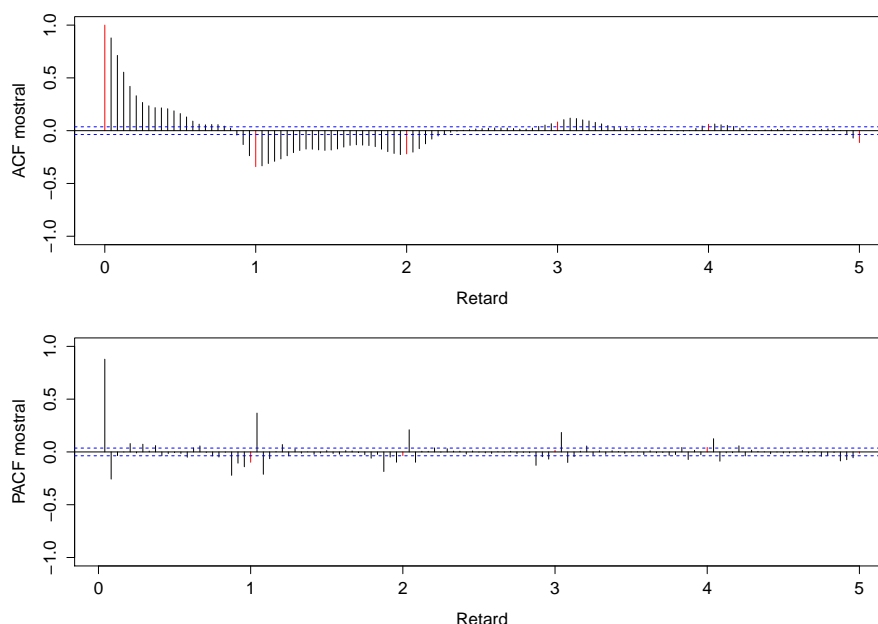


Figura A.5: ACF i PACF mostrals.

Com s'han aplicat diferències d'ordre 24, s'ha de tenir en compte que el model estarà format per una part regular (en negre) i per una d'estacional (en vermell). Analitzant l'ACF es determinarà la part MA del model mentre que amb la PACF la part AR. També es comprova que la sèrie presenta ergodicitat.

Si s'analitza la part regular, es pot observar que l'ACF presenta un decreixement geomètric de manera que no aporta informació sense analitzar abans la PACF. En aquesta última, es pot observar com els coeficients d'autocorrelació són significatius fins el retard número 9. És cert que posteriorment hi ha retards amb valors molt superiors, però s'han de tenir en compte dos fenòmens.

El primer és que un 5 % dels retards podrien semblar significatius per atzar (degut al 95 % de significació). El segon és l'estacionalitat de la sèrie. Valors propers a retards múltiples de 24 semblen molt elevats però realment estan contaminats per l'estacionalitat de la sèrie. Així doncs, es considera que hi ha una part AR d'ordre 9 ($p = 9$)

Pel que fa a la part estacional (en vermell), si s'analitza l'ACF destaquen els dos primers retards (ambdós negatius) mentre que en la PACF hi ha un petit decreixement geomètric. Per aquesta raó es decideix identificar la part estacional del model com una MA d'ordre 2 ($Q = 2$).

Així doncs, es considera que la sèrie de preus es comporta com un $SARIMA(9, 0, 0)(0, 1, 2)_{24}$:

$$(1 - \phi_1 B - \dots - \phi_9 B^9)(1 - B^{24})z_t = (1 + \Theta_{24}B^{24} + \Theta_{48}B^{48})a_t \quad (\text{A.1})$$

Estimació del model

Un cop identificat el model, aquest s'ha d'estimar. Així doncs, es forçarà a la sèrie de preus a comportar-se com un l'equació A.1 i d'aquesta manera, estimar els seus coeficients. És important comprovar que el terme independent del model és pràcticament 0. Si no fos així, s'hauria de tenir en compte en la formulació del model.

```
arima(x = Gener, order = c(9, 0, 0), seasonal = list(order = c(0, 1, 2),
period = 24))

Coefficients:
      ar1      ar2      ar3      ar4      ar5      ar6      ar7      ar8      ar9
  1.1761 -0.2986  0.0079 -0.0802  0.0527 -0.0174  0.0631 -0.0985  0.0940
s.e.  0.0188  0.0288  0.0292  0.0292  0.0293  0.0293  0.0292  0.0287  0.0186
      sma1      sma2
 -0.9597 -0.0403
s.e.   0.0214  0.0188

sigma^2 estimated as 0.0334:  log likelihood = 751.06,  aic = -1478.13
```

Figura A.6: Model univariant.

D'aquest model destaca que la major part dels coeficients semblen significatius i que té una variància de 0.0334. Aquest últim valor és important ja que permetrà veure l'evolució del model.

En aquest cas, amb les ACF i PACF donades es podria considerar que un $SARIMA(9, 0, 0)(1, 1, 1)_{24}$ s'adaptaria correctament. S'ha estimat aquest model i el seu AIC ha sigut de -1478,26 de manera que el primer model compleix millor el criteri de selecció.

Anàlisi dels residus

Tot i que aquest model no serà definitiu, s'utilitzarà com a base per realitzar els següents. Per aquesta raó es comprovarà que els residus es comporten correctament. Si no fos així, s'hauria de tornar a identificar i estimar el model.

En primer lloc es comprova que tinguin mitjana nul·la (sempre s'ha de tenir en compte el possible 5 % de valors atípics). Es pot observar un comportament acceptable a la Figura A.7.

En segon lloc s'ha de comprovar que com a mínim fins el residu número 10 no existeixen residus que expliquen la sèrie. Els estadístics de Ljung-Box indiquen que fins el residu número 18 es compleix la condició d'independència dels residus, i per tant, els resultats fins aquest punt són satisfactoris.

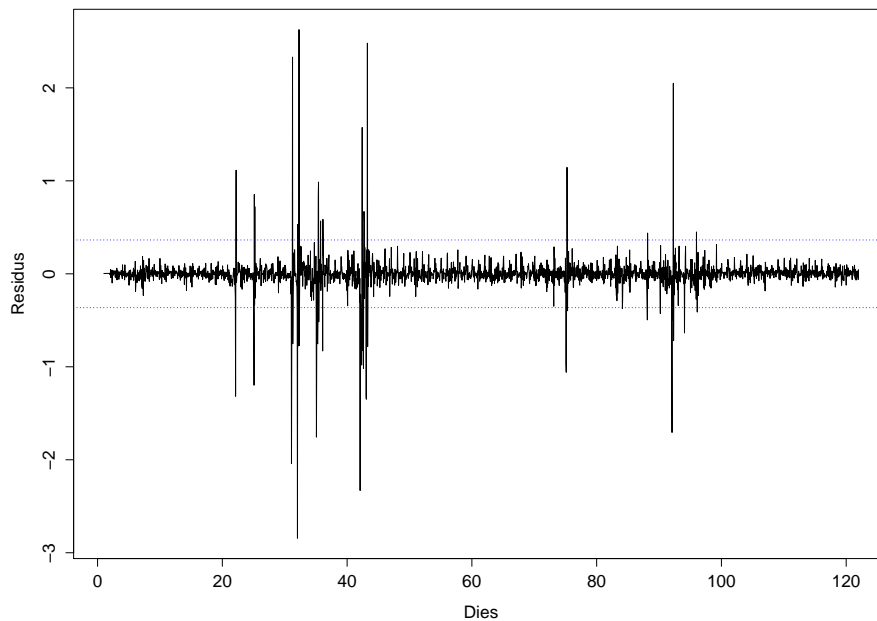


Figura A.7: Representació del comportament dels residus. S'ha de tenir en compte que s'ha transformat la sèrie mitjançant logaritmes

Posteriorment s'analitza si els residus presenten una distribució normal. A l'histograma de la Figura A.8 es pot observar que es comporten amb aquesta normalitat però presenten una particularitat. La densitat de probabilitat al voltant del zero és més elevada fent que la campana s'estiri respecte una distribució normal teòrica. Aquest fet reflecteix l'alta volatilitat de les dades ja que aquesta distribució representa un increment de la densitat de probabilitat a les cues de la campana. Aquesta volatilitat també es pot veure en el diagrama Q-Q que presenta cues pesants:

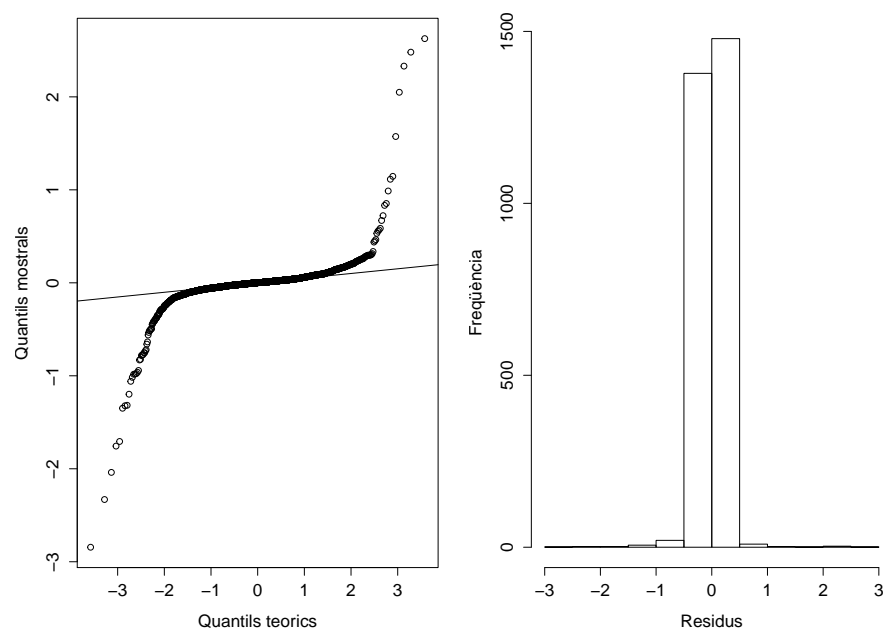


Figura A.8: Histograma i diagrama Q-Q dels residus.

Per últim, es representa l'ACF i la PACF dels residus i dels residus al quadrat per comprovar novament que aquests no expliquen el model. Com es pot observar en la Figura A.9, al voltant del residu 20 sembla que hi ha valors significatius. Tot i així, son molt allunyats i no es considerarà un problema en aquesta fase.

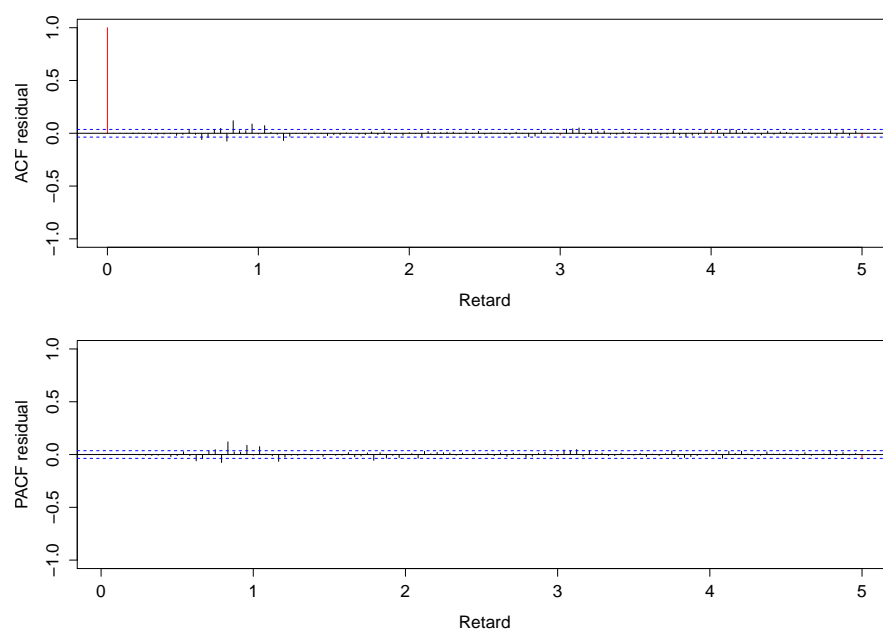


Figura A.9: ACF i PACF dels residus.

A més, s'observen (Figura A.10) les corbes típiques de les series volàtils:

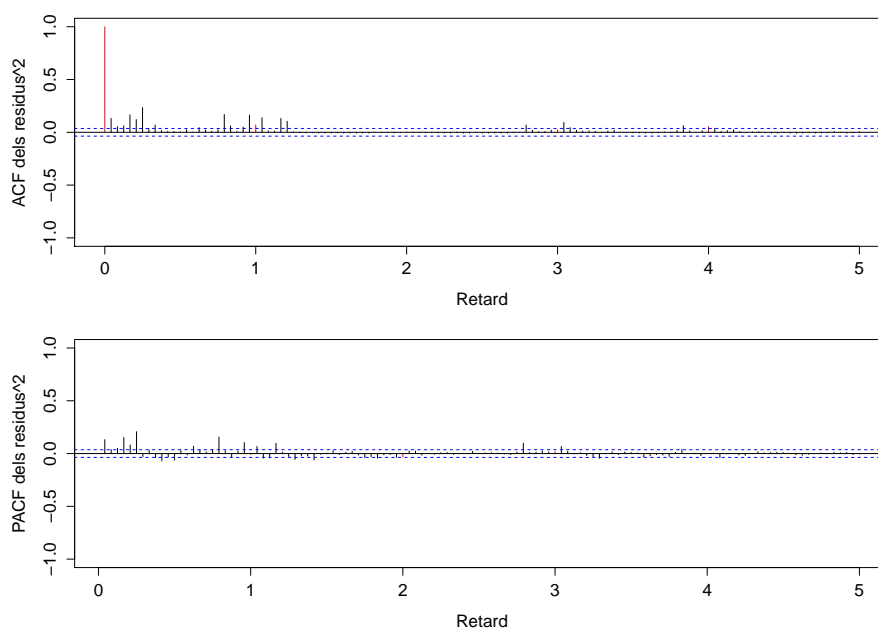


Figura A.10: ACF i PACF dels residus al quadrat.

A.2 Determinació del model amb variables exògenes

Com s'ha comentat al punt 4.2.3, el model multivariant consta de dues parts. En la primera, es determina un model de regressió múltiple mitjançant mínims quadrats mentre que en la segona es determina un model ARMA pels residus del primer (ξ_t). Es suposarà que aquests residus es comporten igual que el model univariant. Si s'introdueixen les variables i el model identificat, R retorna directament l'estimació dels coeficients:

```
arima(x = Gener, order = c(9, 0, 0), seasonal = list(order = c(0, 1, 2),
period = 24), xreg = exo)
```

Coefficients:

	ar1	ar2	ar3	ar4	ar5	ar6	ar7	ar8	ar9
	1.1359	-0.2860	0.0112	-0.0949	0.0488	-0.0147	0.0533	-0.0968	0.0716
s.e.	0.0190	0.0284	0.0287	0.0287	0.0289	0.0289	0.0287	0.0283	0.0187

	sma1	sma2	Preu168	PrevisioDemandaMes	PrevisioEolicaMes
	-0.9713	-0.0287	71.2242	0.2559	-0.4872
s.e.	0.0202	0.0189	9.1310	0.0436	0.0646

	ReservesHidroelectriciquesMes	NuclearDisponibleMes	OMIPMes
	0.6538	-0.2125	46.4154
s.e.	0.6817	0.2596	18.8715

sigma^2 estimated as 0.03172: log likelihood = 830.06, aic = -1624.12

Figura A.11: Model multivariant.

Es pot observar com la variància del model a passat de 0.0334 a 0.03172. És a dir, s'ha millorat. El mateix ha passat amb l'AIC que ha disminuït 146.18 unitats.

El criteri utilitzat per determinar si una variable és significativa pel model és que el quocient entre el coeficient i la desviació estàndard sigui superior a 2. Amb aquest criteri les reserves hidroelèctriques i l'energia nuclear són no significatives. Tot i així, no es descarten del model ja que, en primer lloc, s'ha de fer un anàlisi dels valors atípics (és possible que depurant aquests valors, aquestes variables passin a ser significatives). A més, en tractar-se de previsions a curt termini, encara que una variable sembli no significativa no es descartarà. Això permetrà veure'n l'evolució.

A.3 Detecció de valors atípics i linealització de la sèrie

Detecció de valors atípics

En el cas d'estudi les bandes fora de les quals es considera que el valor és atípic s'han ubicat aproximadament a 3.5 desviacions estàndard (tant en valors positius com en negatius). A la Figura A.12 es representen en vermell els valors dintre de les bandes mentre que la resta són considerats valors atípics:

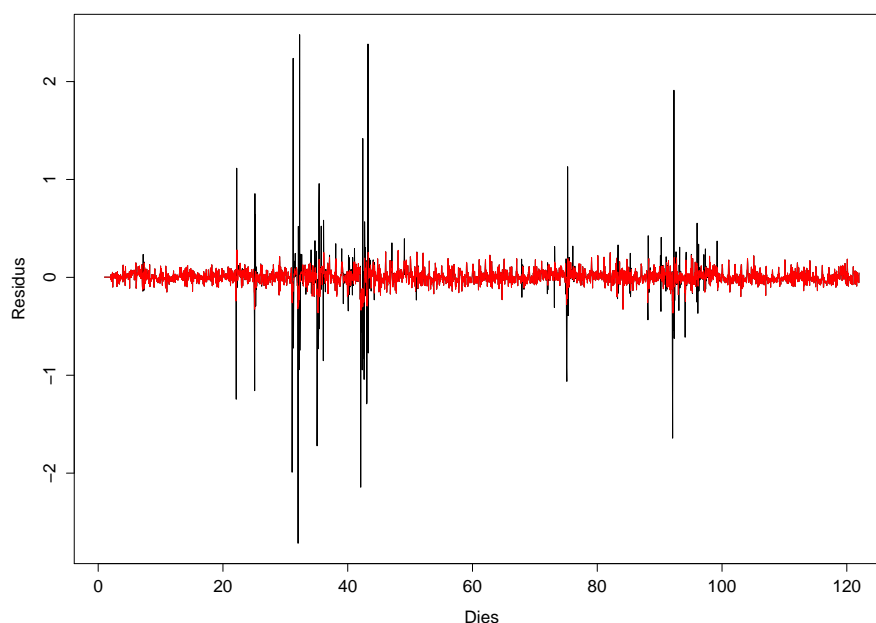


Figura A.12: Representació dels valors observats. En negre els que es consideren atípics.

En aquest cas, s'han determinat 85 valors atípics que representen un 2,95 % dels valors de la sèrie. A la Figura A.13, es pot veure l'observació atípica (*Obs*), el tipus (*type_dected*), el número d'unitats de desviació (*W_coef*) i el grau d'anomalia del valor (com més gran és *ABS_L_Ratio*, més anòmal és el valor). Si es tracta d'una observació més petita del que s'esperava *W_coef* serà negatiu i en cas contrari, positiu.

	Obs	type_detected	W_coeff	ABS_L_Ratio
94	149	A0	-0.1842603	3.856855
13	508	TC	-1.1243397	10.135942
77	509	TC	-0.3036751	4.294200
18	510	TC	0.9622263	9.479470
93	578	TC	-0.2608370	3.857548
8	579	A0	-0.9462129	10.948641
30	582	TC	0.6051426	6.760988
10	723	TC	-1.2786712	10.917363
48	723	TC	-0.4510833	5.635355
11	724	TC	-1.2471600	10.862492
64	724	TC	-0.3500758	4.704271
47	725	TC	-0.4553834	5.658221
85	725	TC	-0.2826888	4.094335
49	726	TC	-0.4633918	5.822820
4	727	TC	2.1574306	15.911251
63	728	TC	0.3345339	4.478390
1	746	TC	-2.4520961	15.894990
72	747	TC	-0.3050980	4.223374
73	748	TC	-0.3529747	4.906326
74	749	TC	-0.4305952	6.022507
39	750	TC	-0.4994515	5.924742
75	750	TC	-0.3710834	5.214388
40	751	TC	-0.5306127	6.337773
2	752	TC	2.4351160	16.509042
95	752	A0	-0.1839412	3.860040
82	796	TC	0.2874359	4.127187
61	811	TC	0.3588786	4.771357
79	817	TC	0.3009524	4.282687
14	819	TC	-1.0273201	9.401184
35	819	TC	-0.5342046	6.173829
15	820	TC	-1.3236997	12.431550
88	820	TC	-0.2650338	3.869048
46	821	TC	-0.4651054	5.747425
45	822	TC	-0.4564202	5.608292
41	823	TC	-0.4902153	5.890128
33	824	TC	-0.5529657	6.305629
22	825	A0	-0.5455260	8.016118
17	827	TC	1.0136914	9.835470
53	835	TC	0.3871755	4.950444
19	843	TC	-0.9569570	9.575226
102	891	TC	0.2414023	3.648145
109	914	TC	0.2316976	3.555921
100	920	TC	-0.2456035	3.694536
52	939	TC	-0.3916089	4.986135
104	942	A0	0.1681463	3.603325
98	955	A0	-0.1753949	3.707263
110	963	TC	0.2307242	3.548652
50	986	A0	0.2849447	5.077087
6	987	TC	-1.5718316	12.284282
24	988	TC	-0.7030490	7.464749
68	989	TC	-0.3790193	5.178216
66	990	TC	-0.3297784	4.462211
90	990	TC	-0.2703290	3.967261
67	991	TC	-0.3994097	5.431774
80	992	TC	-0.2951318	4.212670
28	993	TC	-0.6426977	7.065009
89	993	TC	-0.2649658	3.878057
37	994	TC	-0.5544552	6.495443

7	995	TC	1.5919361	12.786852
69	997	A0	0.2340252	4.529604
20	1000	TC	-0.9017271	9.151788
36	1001	A0	-0.3714125	6.098411
96	1003	TC	0.2519932	3.754944
84	1005	TC	0.2830081	4.087180
16	1010	TC	-1.0247066	9.780776
12	1011	TC	-1.1332418	10.040139
51	1011	TC	-0.4043490	5.126449
55	1012	TC	-0.4426684	5.715597
54	1013	TC	-0.3847973	4.940671
34	1014	TC	-0.5521808	6.340108
3	1015	TC	2.2665195	16.031584
106	1018	A0	0.1659864	3.572830
91	1107	TC	0.2642230	3.887730
76	1155	TC	0.3037737	4.282020
99	1201	TC	-0.2464765	3.698986
87	1606	A0	0.1908416	3.922919
111	1705	A0	-0.1617283	3.519089
60	1732	TC	-0.3614503	4.786822
108	1776	A0	-0.1647556	3.561863
27	1779	TC	-0.6460320	7.041406
23	1780	TC	-0.7417518	7.801193
38	1780	TC	-0.5229415	6.166236
25	1781	TC	-0.6910701	7.406540
21	1782	A0	-0.6317919	9.182695
62	1783	TC	0.3354904	4.475765
29	1784	TC	0.6333368	7.020947
97	1803	TC	0.2507485	3.745410
107	1975	TC	-0.2336114	3.569725
65	1977	TC	0.3342665	4.507513
78	2023	A0	-0.2147253	4.300159
32	2092	A0	-0.3939990	6.299736
103	2140	TC	0.2376468	3.599393
44	2142	A0	-0.3301060	5.695475
105	2161	A0	-0.1679980	3.608208
26	2186	TC	-0.6479142	7.002396
9	2187	TC	-1.3249793	11.087500
59	2187	TC	-0.4487861	5.920132
58	2188	TC	-0.4983004	6.533986
57	2189	TC	-0.3711874	4.831824
43	2190	TC	-0.4714728	5.730217
5	2192	TC	1.8368697	13.996571
56	2211	TC	-0.3787926	4.911121
86	2214	TC	0.2720255	3.950465
42	2235	TC	-0.4831758	5.839532
92	2239	TC	0.2611501	3.852321
31	2280	A0	-0.3987012	6.331800
83	2284	TC	-0.2861791	4.121138
71	2287	TC	0.3096863	4.273783
101	2307	A0	0.1730771	3.684066
81	2313	TC	0.2918537	4.178384
70	2358	TC	0.3240926	4.458596

Figura A.13: Valors atípics.

Un cop identificats aquests valors s'han d'analitzar i comprovar que són atípics. A la Figura A.14, es representen els valors de les observacions compreses entre la 500 i la 512 per poder analitzar per què els valors 508, 509 i 510 són atípics:

	exp(Gener[500:512]) - 1		Preu168	PrevisioDemandaMes	PrevisioEolicaMes
500	38.97	0.005601		2.5215	0.9317
501	50.42	0.006711		2.8178	0.8761
502	52.70	0.006988		2.8435	0.8330
503	53.03	0.006975		2.7113	0.8349
504	45.01	0.006019		2.5111	0.8389
505	45.91	0.006019		2.2451	0.8970
506	35.46	0.005523		2.0989	0.8619
507	24.00	0.005303		2.0136	0.8908
508	5.00	0.004338		1.9990	0.9181
509	5.00	0.004237		1.9684	0.9260
510	24.00	0.004570		2.0284	0.9977
511	40.01	0.005665		2.2658	1.0202
512	55.26	0.006600		2.6643	1.0925
	ReservesHidroelectricquesMes		NuclearDisponibleMes	OMIPMes	
500	1.0379		0.66818	0.004593	
501	1.0379		0.66818	0.004593	
502	1.0379		0.66818	0.004593	
503	1.0379		0.66818	0.004593	
504	1.0379		0.66818	0.004593	
505	1.0379		0.56690	0.004593	
506	1.0379		0.56690	0.004593	
507	1.0379		0.56690	0.004593	
508	1.0379		0.56690	0.004593	
509	1.0379		0.56690	0.004593	
510	1.0379		0.56690	0.004593	
511	1.0379		0.56690	0.004593	
512	1.0379		0.56690	0.004593	

Figura A.14: Valors compresos entre les observacions 500 i 512.

Es pot observar una baixada brusca en el preu de l'observació 508 (un salt de 20 unitats mentre que en la resta d'observacions no sobrepassa les 10 unitats). Posteriorment la observació 509 manté el mateix valor, raó per la qual és un altre atípic. Curiosament, la 510 torna a ser 24 però com la sèrie ja considerava que s'havia d'estar en l'entorn del 5, apareix com una altra anomalia. En aquest cas es pot veure que és un TC ja que hi ha un canvi i un retorn gradual al valor original. Per exemple, la següent observació (579) es tracta d'un AO tal i com s'observa a la Figura A.15:

	exp(Gener[575:585]) - 1		Preu168	PrevisioDemandaMes	PrevisioEolicaMes
575	53.06	0.005523		3.0412	1.0611
576	50.00	0.005303		2.7584	1.0324
577	40.00	0.005306		2.5410	1.0811
578	22.07	0.004800		2.3768	1.0637
579	4.90	0.004200		2.2701	1.0495
580	10.07	0.003701		2.2258	1.0512
581	10.07	0.003491		2.2061	1.0635
582	30.00	0.004166		2.2622	0.9581
583	40.55	0.005000		2.4705	0.8635
584	55.13	0.005613		2.8911	0.7480
585	57.13	0.006019		3.1564	0.7115
	ReservesHidroelectriciquesMes	NuclearDisponibleMes	OMIPMes		
575	1.0156	0.56612	0.005675		
576	1.0156	0.56612	0.005675		
577	1.0156	0.56705	0.005675		
578	1.0156	0.56705	0.005675		
579	1.0156	0.56705	0.005675		
580	1.0156	0.56705	0.005675		
581	1.0156	0.56705	0.005675		
582	1.0156	0.56705	0.005675		
583	1.0156	0.56705	0.005675		
584	1.0156	0.56705	0.005675		
585	1.0156	0.56705	0.005675		

Figura A.15: Valors compresos entre les observacions 575 i 585.

Linealització de la sèrie

Un cop analitzades les observacions atípiques es linealitzarà la sèrie de preus. És a dir, es substituiran els valors atípics pels valors que el model esperava trobar. A la Figura A.16, es poden veure els nous valors linealitzats en un dels casos exposats a l'apartat anterior:

	exp(Gener.lin[500:512]) - 1		Preu168	PrevisioDemandaMes	PrevisioEolicaMes
500	65.45	0.005601		2.5215	0.9317
501	64.50	0.006711		2.8178	0.8761
502	66.01	0.006988		2.8435	0.8330
503	60.24	0.006975		2.7113	0.8349
504	58.49	0.006019		2.5111	0.8389
505	53.32	0.006019		2.2451	0.8970
506	41.85	0.005523		2.0989	0.8619
507	3.00	0.005303		2.0136	0.8908
508	1.00	0.004338		1.9990	0.9181
509	3.00	0.004237		1.9684	0.9260
510	25.33	0.004570		2.0284	0.9977
511	36.57	0.005665		2.2658	1.0202
512	56.03	0.006600		2.6643	1.0925
	ReservesHidroelectriciquesMes	NuclearDisponibleMes		OMIPMes	
500	1.0379		0.66818	0.004593	
501	1.0379		0.66818	0.004593	
502	1.0379		0.66818	0.004593	
503	1.0379		0.66818	0.004593	
504	1.0379		0.66818	0.004593	
505	1.0379		0.56690	0.004593	
506	1.0379		0.56690	0.004593	
507	1.0379		0.56690	0.004593	
508	1.0379		0.56690	0.004593	
509	1.0379		0.56690	0.004593	
510	1.0379		0.56690	0.004593	
511	1.0379		0.56690	0.004593	
512	1.0379		0.56690	0.004593	

Figura A.16: Valors linealitzats compresos entre les observacions 500 i 512.

A la Figura A.17 s'ha representat la diferència entre la sèrie original i la linealitzada. Es pot observar com la majoria de les correccions s'han fet perquè el valor de la sèrie original era molt més baix del que s'esperava.

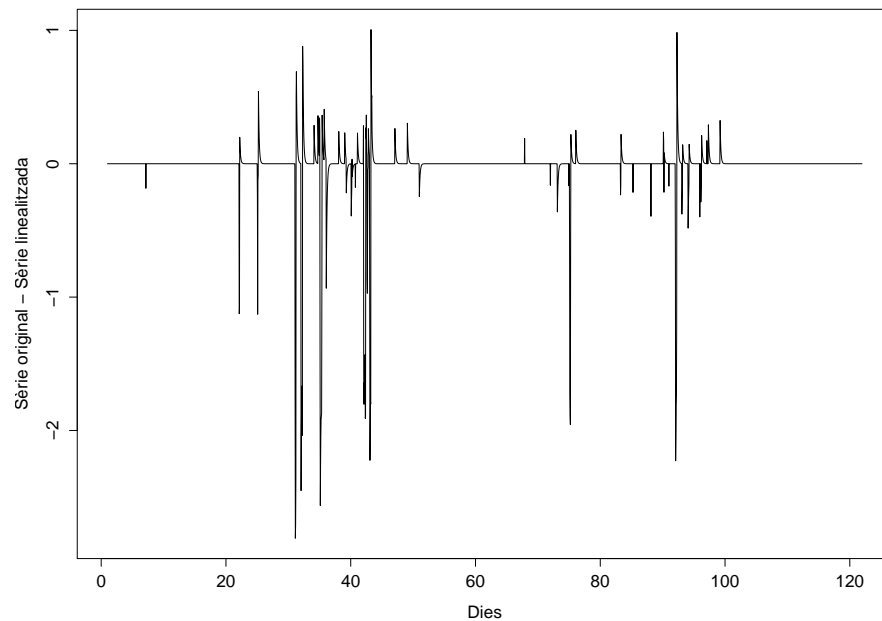


Figura A.17: Representació de la sèrie original menys la linealitzada.

A.4 Determinació del model final

Identificació i estimació del model final

Amb aquesta nova sèrie linealitzada es tornarà a repetir el procediment per obtenir un model multivariant definitiu. És a dir, es tornarà a analitzar l'ACF i la PACF mostrals i posteriorment s'estimaran els coeficients del model identificat.

En aquest cas, es tornen a aplicar logaritmes a la sèrie i s'apliquen diferències per eliminar l'estacionalitat. A la Figura A.18 es pot veure que l'ACF i la PACF resultant és molt semblant al que s'havia obtingut amb la sèrie original. Així doncs, el model teòric continuarà essent el mateix.

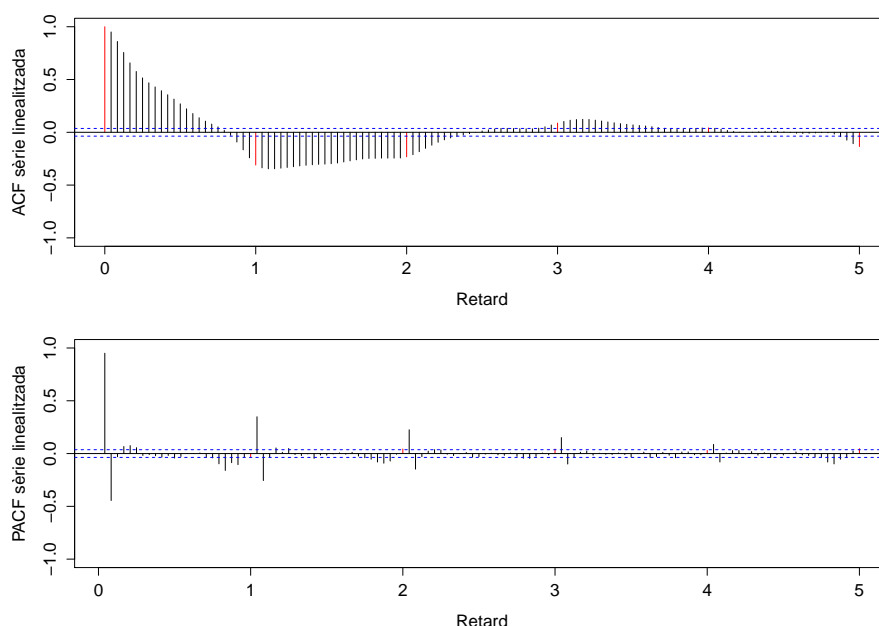


Figura A.18: ACF i PACF mostrals de la sèrie linealitzada.

Així doncs, es torna a estimar el model. A la Figura A.19, es pot observar com variables que abans no eren significatives, ara ho són. Això es degut a la depuració realitzada per l'anàlisi de valors atípics. A més, s'observa com la variància ha disminuït considerablement (en el model anterior era de 0.03172, fet que reflecteix una millora del model).

```
arima(x = Gener.lin, order = c(9, 0, 0), seasonal = list(order = c(0, 1, 2),
period = 24), xreg = exo)
```

Coefficients:

	ar1	ar2	ar3	ar4	ar5	ar6	ar7	ar8	ar9
	1.3357	-0.3594	-0.0803	-0.0505	0.0279	0.0559	-0.0141	-0.0244	0.0257
s.e.	0.0190	0.0313	0.0320	0.0321	0.0321	0.0320	0.0320	0.0313	0.0188

	sma1	sma2	Preu168	PrevisioDemandaMes	PrevisioEolicaMes
	-0.8650	-0.1104	39.5927	0.2422	-0.3351
s.e.	0.0185	0.0184	3.3846	0.0197	0.0375

	ReservesHidroelectriciquesMes	NuclearDisponibleMes	OMIPMes
	0.8244	-0.2768	14.8585
s.e.	0.5058	0.1039	8.2228

sigma^2 estimated as 0.004132: log likelihood = 3761.67, aic = -7487.35

Figura A.19: Model multivariant final.

Pel que fa l'AIC, també a disminuït considerablement però s'ha de tenir en compte que la sèrie linealitzada incorpora a paràmetres (equació 4.35). Per tant, el valor real d'aquest coeficient serà:

$$AIC_c = AIC + 2 \cdot 85 = -7317.35 \quad (A.2)$$

Tot i així, la diferència continua essent gran.

Anàlisi dels residus del model final

Es tornen a analitzar els residus tal i com s'ha fet anteriorment amb el model univariant. En aquest cas es serà més exigent atès que es tracta del model final. Es pot observar (Figura A.20) com els valors atípics continuen dintre del 5 % que s'accepta (lògicament, el percentatge disminueix degut a la linealització de la sèrie). A més, els estadístics de Ljung-Box han passat del residu número 18 del primer model al número 22. Es considera que aquest resultat és satisfactori.

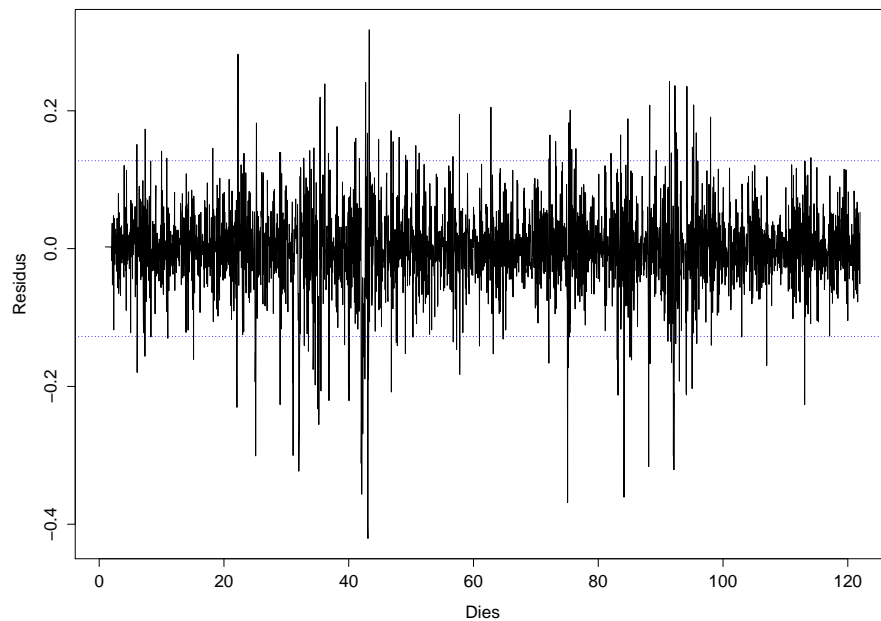


Figura A.20: Representació del comportament dels residus de la sèrie linealitzada.

Pel que fa a la normalitat dels residus, es pot observar com la volatilitat del primer model ha disminuït. Així doncs, l'histograma (Figura A.21) presenta una campana més semblant a la teòrica i tot hi haver cues pesants, aquestes son lleugerament més discretes. La resta de gràfics mostren el mateix comportament de manera que els residus del model es comporten correctament.

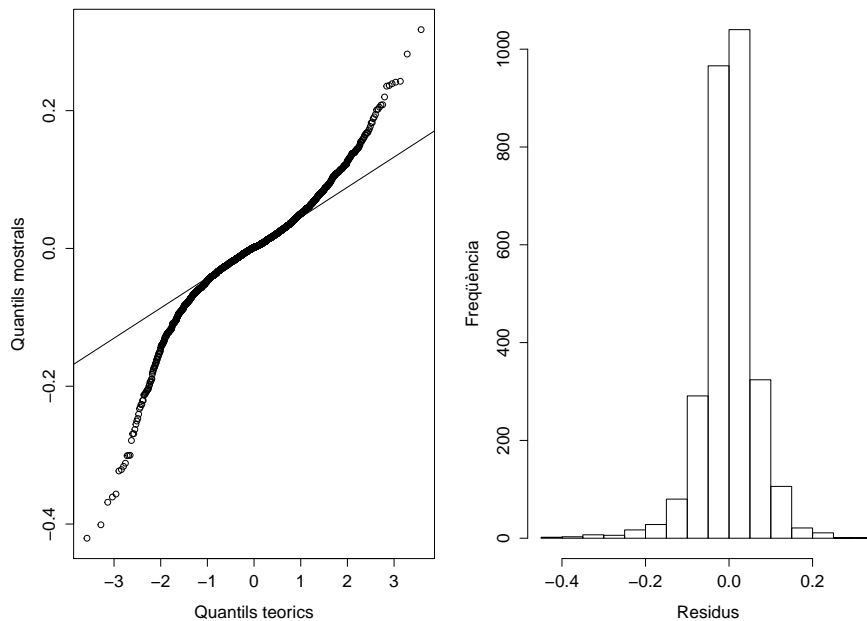


Figura A.21: Histograma i gràfica Q-Q dels residus.

Validació i actualització del model final

Un cop s'ha comprovat que els residus es comporten correctament, s'ha de decidir si el model generat (en aquest cas $M2$) serà el model utilitzat per predir el període 2 (o en el seu defecte es continuarà amb $M1$).

D'aquesta manera es mesura la quantitat de valors atípics que hi ha en l'últim mes de generació.

En la generació de $M2$ hi ha 111 valors atípics, però només 16 d'ells estan presents en l'últim mes. Per tant, en estar molt per sota del 40% que s'ha establert com a límit, s'actualitza el model i s'utilitza $M2$.

Annex B

Codi R

B.1 Sèries temporals

B.1.1 Codi general

```
## Calendari

library(chron)

cal = dates("12/31/2005") + 1:nrow(Preus)
cal168 = dates("12/31/2005") + 8:nrow(Preus)
caleolica = dates("12/31/2010") + 1:nrow(PrevisioEolica)
caldemanda = dates("12/31/2009") + 1:nrow(PrevisioDemanda)
calhidroelectriques = dates("10/04/2010") + 2:nrow(ReservesHidroelectriques)
calnuclear = dates("05/10/2010") + 1:nrow(NuclearDisponible)
calomip = dates("05/22/2011") + 1:nrow(OMIP2D)

## Introducció de les dades

datafin = "1/3/2012"

M1 = ts(c(t(Preus[cal >= dates(datafin) - 120 & cal <= dates(datafin), ])),
        freq = 24)

Preu168 = ts(c(t(Preus[cal168 >=
dates(datafin) - 120 & cal168 <= dates(datafin),
])), freq = 24)

PrevisioDemanda = ts(c(t(PrevisioDemanda[caldemanda >= dates(datafin) - 120 &
caldemanda <= dates(datafin), ])), freq = 24)

PrevisioEolica = ts(c(t(PrevisioEolica[caleolica >= dates(datafin) - 120 &
caleolica <= dates(datafin), ])), freq = 24)
```

```

ReservesHidroelectriques = ts(c(t(ReservesHidroelectriques
[calhidroelectriques >=dates(datafin) - 120 & calhidroelectriques <= dates
(datafin), ])), freq = 24)

NuclearDisponible = ts(c(t(NuclearDisponible[calnuclear >= dates(datafin) -
120 & calnuclear <= dates(datafin), ])), freq = 24)

OMIP2D = ts(c(t(OMIP2D[calomip >= dates(datafin) - 120 & calomip <= dates(datafin),
])), freq = 24)

## Transformacions de la sèrie original

M1 = log(M1 + 1)

diffM1 = diff(M1, 24)

## Anàlisi de l'ACF i la PACF. Identificació del model

ACFPACF(diffM1)

## Estimació del model univariant. SARIMA(9,0,0)(0,1,2)24

(ModelM10 = arima(diffM1, order = c(9, 0, 0), seasonal = list(order = c(0, 0,
2), period = 24)))

(ModelM11 = arima(M1, order = c(9, 0, 0), seasonal = list(order = c(0, 1, 2),
period = 24)))

Bondat(ModelM11)

## Estimació del model multivariant. SARIMAX(9,0,0)(0,1,2)24

exo = data.frame(Preu168, PrevisioDemandaMes, PrevisioEolicaMes,
ReservesHidroelectriquesMes, NuclearDisponibleMes, OMIPMes)/10000

(ModelM12 = arima(M1, order = c(9, 0, 0), seasonal = list(order =
c(0, 1, 2),
period = 24), xreg = exo))

Bondat(ModelM12)

## Detecció d'atípics i linealització de la sèrie

mod.atip = outdetec(ModelM12, dif = c(0, 24), crit = 3.5, LS = F)

mod.atip$sigma2

mod.atip$atip

```

```

mod.atip$atip[order(mod.atip$atip[, 1
]), ]

cbind(exp(M1[500:512]) - 1, exo[500:512, ])

M1.lin = lineal(M1, mod.atip$atip)

## Determinació del model final

diffM1.lin = diff(M1.lin, 24)

ACFPACF(diffM1.lin)

(ModelM12.lin = arima(M1.lin, order = c(9, 0, 0), seasonal = list(order = c(0,
1, 2), period = 24), xreg = exo))

Bondat(ModelM12.lin)

## Validació. Realització d'una predicció

exo.nou = data.frame(Preu168 = ts(c(t(Preus[cal168 == dates(datafin) + 1, ])),
freq = 24), PrevisioDemandaMes = ts(c(t(PrevisioDemanda[caldemanda ==
dates(datafin) + 1, ])), freq = 24),

PrevisioEolicaMes = ts(c(t(PrevisioEolica[caleolica ==
dates(datafin) + 1, ])), freq = 24), ReservesHidroelectriquesMes =
ts(c(t (ReservesHidroelectriques[calhidroelectriques ==
dates(datafin) + 1, ])), freq = 24), NuclearDisponibleMes =
ts(c(t(NuclearDisponible[calnuclear ==
dates(datafin) + 1, ])), freq = 24), OMIPMes = ts(c(t(OMIP2D[calomip ==
dates(datafin) + 1, ])), freq = 24))/10000

pr = predict(ModelM12.lin, n.ahead = 24, newxreg = exo.nou)

linf = exp(pr$pred - 1.96 * pr$se) - 1

pred = exp(pr$pred) - 1

lsup = exp(pr$pred + 1.96 * pr$se) - 1

predic = ts.union(linf, pred, lsup)

M1.obs = ts(c(t(Preus[cal >= dates(datafin) - 120 & cal <= dates(datafin) +
1, ])), freq = 24)

ts.plot(window(M1.obs, start = 116), linf, pred, lsup, col = c(1, 4, 2, 4),
type = "o")

obs = window(M1.obs, start = 122)

```

```

ts.union(pred, obs, err = obs - pred, err.rel = (obs - pred)/obs)

(MAPE = sum(abs(obs - pred)/obs)/24)

## Funció que permet fer les prediccions a 120 dies.

serie.ori = M1.lin

ndays = 120

MAPEReal = rep(0, ndays)
MAPECorregit = rep(0, ndays)

predicTot = NULL
dataeval = NULL
MAPEReal = NULL
for (k in 1:ndays) {
  (dataeval[[k]] = dates(datafin) + k)
  M1.lin = c(tail(serie.ori, length(M1.lin) - 24 * (k - 1)), log(c(t(Preus[cal >
    dates(datafin) & cal <= dates(datafin) + k - 1, ])) + 1))

  exo.nou = data.frame(Preu168 = ts(c(t(Preus[cal168 == dates(datafin) + k,
    ])), freq = 24), PrevisioDemandaMes = ts(c(t(PrevisioDemanda[caldemanda
    == dates(datafin) + k, ])), freq = 24), PrevisioEolicaMes =
    ts(c(t(PrevisioEolica[caleolica == dates(datafin) + k, ])), freq = 24),
  ReservesHidroelectriquesMes = ts(c(t(ReservesHidroelectriques[calhidroelectriques
    == dates(datafin) + k, ])), freq = 24), NuclearDisponibleMes =
    ts(c(t(NuclearDisponible[calnuclear ==
    dates(datafin) + k, ])), freq = 24), OMIPMes = ts(c(t(OMIP2D[calomip ==
    dates(datafin) + k, ])), freq = 24))/10000

  pr = predict(ModelM12.lin, n.ahead = 24, newxreg = exo.nou)
  linf = exp(pr$pred - 1.96 * pr$se) - 1
  pred = exp(pr$pred) - 1
  lsup = exp(pr$pred + 1.96 * pr$se) - 1

  M1.obs = ts(c(t(Preus[cal >=
    dates(datafin) - 120 + k - 1 & cal <= dates(datafin) +
    k, ])), freq = 24)

  ts.plot(window(M1.obs, start = 116), linf, pred, lsup, col = c(1, 4, 2,
    4), type = "o")

  obs = window(M1.obs, start = 122)
  predicTot[[k]] = ts.union(pred, obs, err = obs - pred, err.rel = (obs -
    pred)/obs)
  mapehores = abs(obs - pred)/obs
  (MAPEReal[k] = mean(mapehores))
  (MAPECorregit[k] = mean(mapehores[mapehores < 1]))
}

```

B.1.2 Funcions auxiliars

```
## ACF i PACF
ACFPACF = function(dif) {
  par(mfrow = c(1, 2))
  acf(dif, ylim = c(-1, 1), lag.max = 5 * 24, lwd = 1, col =
    c(2, rep(1, 23)))
  pacf(dif, ylim = c(-1, 1), lag.max = 5 * 24, lwd = 1, col
    = c(rep(1, 23),
      2))
}

## Bondat d'ajust
Bondat = function(mod) {
  par(mar = c(3, 1, 1, 1))
  tsdiag(mod, gof.lag = 5 * 24)

  ## Variància dels residus constant
  plot(resid(mod))
  abline(h = 2 * sd(resid(mod)), lty = 3, col = 4)
  abline(h = -2 * sd(resid(mod)), lty = 3, col = 4)

  ### Residus normals
  qqnorm(resid(mod))
  qqline(resid(mod))
  hist(resid(mod))

  ### Independència dels residus
  par(mfrow = c(1, 2))
  acf(resid(mod), ylim = c(-1, 1), lag.max = 5 * 24, lwd = 1,
    col = c(2, rep(1, 23)))
  pacf(resid(mod), ylim = c(-1, 1), lag.max = 5 * 24, lwd = 1,
    col = c(rep(1, 23), 2))

  ### Independència dels residus^2 (volatilitat)
  par(mfrow = c(1, 2))
  acf(resid(mod)^2, ylim = c(-1, 1), lag.max = 5 * 24, lwd = 1,
    col = c(2, rep(1, 23)))
  pacf(resid(mod)^2, ylim = c(-1, 1), lag.max = 5 * 24, lwd = 1,
    col = c(rep(1, 23), 2))
}

outdetec <- function(object, dif = c(0, 0), crit, LS = T) {
  residuals <- object$residuals
  m <- length(residuals)

  piweight <- -ARMAtoMA(ar = -object$model$theta, ma = -object$model$phi,
    lag.max = m + sum(dif))
}
```

```

if (dif[1] != 0)
  for (i in 1:dif[1]) piweight <- c(piweight, 0) - c(-1, piweight)
if (length(dif) > 1) {
  for (i in 2:length(dif)) {
    if (dif[i] > 1)
      piweight <- c(piweight, rep(0, dif[i])) - c(rep(0, dif[i] -
        1), -1, piweight)
  }
}
piweight <- piweight[1:m]

atip <- NULL

num <- NULL
type <- NULL
wcoeff <- NULL
LCrit <- NULL
if (crit <= 0) {
  cat("The Critical value may be positive")
}

va <- mean(residuals^2)

c <- cumsum(piweight) - 1

d <- rep(0, m)
delta <- 0.7
d[1] <- piweight[1] - delta
for (i in 2:m) d[i] <- delta * d[i - 1] + piweight[i]

sum1 <- 1 + sum(piweight * piweight)
sum2 <- 1 + sum(c * c)
sum3 <- 1 + sum(d * d)

maxL <- crit + 1

while (maxL > crit) {
  ka1 <- sum1
  ks1 <- sum2
  kt1 <- sum3
  maxL <- 0
  for (i in 1:m) {

    suma1 <- sum(residuals[i:m] * c(1, -piweight[1:(m - i)]))
    suma2 <- sum(residuals[i:m] * c(1, -c[1:(m - i)]))
    suma3 <- sum(residuals[i:m] * c(1, -d[1:(m - i)]))

    ka1 <- ka1 - piweight[m - i + 1] * piweight[m - i + 1]
    w_ao <- suma1/ka1
    v_ao <- va/ka1
    l_ao <- w_ao/sqrt(v_ao)
  }
}

```



```

ks1 <- ks1 - c[m - i + 1] * c[m - i + 1]
w_ls <- suma2/ks1
v_ls <- va/ks1
l_ls <- w_ls/sqrt(v_ls)

kt1 <- kt1 - d[m - i + 1] * d[m - i + 1]
w_tc <- suma3/kt1
v_tc <- va/kt1
l_tc <- w_tc/sqrt(v_tc)

if (abs(l_ao) > maxL & i > sum(dif)) {
  maxL <- abs(l_ao)
  t <- i
  w <- w_ao
  v <- v_ao
  ts <- "A0"
}
if (abs(l_ls) > maxL & LS == T & i != m & i > sum(dif)) {
  maxL <- abs(l_ls)
  t <- i
  w <- w_ls
  v <- v_ls
  ts <- "LS"
}
if (abs(l_tc) > maxL & i != m & i > sum(dif)) {
  maxL <- abs(l_tc)
  t <- i
  w <- w_tc
  v <- v_tc
  ts <- "TC"
}
}

if (maxL > crit) {
  if (ts == "A0")
    residuals[t:m] <- residuals[t:m] + w * c(-1, piweight[1:(m -
      t)])
  if (ts == "LS")
    residuals[t:m] <- residuals[t:m] + w * c(-1, c[1:(m - t)])
  if (ts == "TC")
    residuals[t:m] <- residuals[t:m] + w * c(-1, d[1:(m - t)])

  val <- mean(residuals^2)
  l <- w/sqrt(v * val/va)
  va <- val

  num <- c(num, t)
  type <- c(type, ts)
  wcoeff <- c(wcoeff, w)
  LCrit <- c(LCrit, abs(l))
}

```

```

        atip <- data.frame(Obs = num, type_detected = type, W_coef
        = wcoeff, ABS_L_Ratio = LCrit)
    }
}
return(list(atip = atip, sigma2 = va, resid = residuals))
}

##Funció per linealitzar la sèrie.

lineal <- function(serie, atip) {

  m <- length(serie)
  for (i in 1:nrow(atip)) {
    t <- atip[i, 1]
    ts <- atip[i, 2]
    w <- atip[i, 3]
    if (ts == "TC")
      serie[t:m] <- serie[t:m] - w * c(1, 0.7^(1:(m - t)))
    if (ts == "LS")
      serie[t:m] <- serie[t:m] - w
    if (ts == "A0")
      serie[t] <- serie[t] - w
  }

  return(serie)
}

```

B.2 Xarxes neuronals

```

## Calendari

library(chron)

cal = dates("12/31/2005") + 1:nrow(Preus)
cal168 = dates("12/31/2005") + 8:nrow(Preus)
caleolica = dates("12/31/2010") + 1:nrow(PrevisioEolica)
caldemanda = dates("12/31/2009") + 1:nrow(PrevisioDemanda)
calhidroelectriques = dates("10/04/2010") + 2:nrow(ReservesHidroelectriques)
calnuclear = dates("05/10/2010") + 1:nrow(NuclearDisponible)
calomip = dates("05/22/2011") + 1:nrow(OMIP2D)

## Selecció de les dades

datafin = "1/31/2012"
umb = 180
Serie = Preus[cal >= dates(datafin) - umb & cal <= dates(datafin), ]

Preu168 = Preus[cal168 >= dates(datafin) - umb & cal168 <= dates(datafin), ]

```

```

PrevisioDemandaMes = PrevisioDemanda[caldemanda >= dates(datafin) - umb &
caldemanda <=dates(datafin), ]

PrevisioEolicaMes = PrevisioEolica[caleolica >= dates(datafin) - umb &
caleolica <=dates(datafin), ]

ReservesHidroelectriquesMes = ReservesHidroelectriques[calhidroelectriques
>=dates(datafin) - umb & calhidroelectriques <= dates(datafin), 1, drop = F]

NuclearDisponibleMes = NuclearDisponible[calnuclear >= dates(datafin) - umb &
calnuclear <= dates(datafin), 1, drop = F]

OMIPMes = OMIP2D[calomip >= dates(datafin) - umb & calomip <= dates(datafin),
1, drop = F]

library(neuralnet)

nlen = nrow(Serie)

PreusI = Serie[-nlen, ]

Preus0 = Serie[-1, ]

data = cbind(PreusI, Preus0, Preu168[-nlen, ], PrevisioDemandaMes[-nlen, ],
PrevisioEolicaMes[-nlen, ], ReservesHidroelectriquesMes[-nlen, ],
NuclearDisponibleMes[-nlen, ], OMIPMes[-nlen, ])

dimnames(data)[[2]] = c(
paste("I", 0:23, sep = ""), paste("0", 0:23, sep = ""),
paste("Preu168", 0:23, sep = "_"), paste("PrevDemMes", 0:23, sep = "_"),
paste("PrevEol", 0:23, sep = "_"), "ResHidro", "NuclDisp", "OMIP")

form = paste(paste(paste("0", 0:23, sep = ""), collapse = "+"), "~",
paste(colnames(data)[-c(25:48)],
collapse = "+"), sep = "")

set.seed(123456)
idx = sample(1:nrow(data), 15, replace = F)

datatrain = data[-idx, ]

dataset = data[idx, ]

(nn = neuralnet(form, datatrain, hidden = 24, rep = 20, err.fct = "sse",
stepmax = 4e+05, threshold = 0.01, lifesign = "full"))

mape = rep(NA, length(idx))

```

```

par(mfrow = c(4, 4), mar = c(3, 3, 1, 1))
for (k in 1:length(idx)) {
  num = idx[k]
  pr = c(unlist(compute(nn, data[num, -c(25:48)], rep = 12)$net.result))
  inp = unlist(data[num, 1:24])
  obs = unlist(data[num, 24 + 1:24])
  plot(obs, type = "o")
  lines(pr, col = 2, type = "o")
  lines(inp, col = 3, type = "o")
  mape[k] = mean(abs(obs - pr)/obs)
}

## Prediccions

ndays = 120

MAPEReal = rep(0, ndays)
MAPECorregit = rep(0, ndays)

predicTot = NULL
dataeval = NULL
MAPEReal = NULL

for (k in 1:ndays) {

  PreusI = Preus[cal == dates(datafin) + k, ]
  Preus0 = Preus[cal == dates(datafin) + k + 1, ]

  Preu168 = Preus[cal168 == dates(datafin) + k, ]
  PrevisioDemandaMes = PrevisioDemanda[caldemanda == dates(datafin) + k, ]
  PrevisioEolicaMes = PrevisioEolica[caleolica == dates(datafin) + k, ]

  ReservesHidroelectriquesMes = ReservesHidroelectriques[calhidroelectriques ==
    dates(datafin) + k, 1, drop = F]
  NuclearDisponibleMes = NuclearDisponible[calnuclear == dates(datafin) +
    k, 1, drop = F]
  OMIPMes = OMIP2D[calomip == dates(datafin) + k, 1, drop = F]

  data.nou = cbind(PreusI, Preu168[-nlen, ], PrevisioDemandaMes[-nlen, ],
    PrevisioEolicaMes[-nlen, ], ReservesHidroelectriquesMes[-nlen, ],
    NuclearDisponibleMes[-nlen, ], OMIPMes[-nlen, ])
  dimnames(data)[[2]] = c(paste("I", 0:23, sep = ""), paste("0", 0:23, sep = ""),
    paste("Preu168", 0:23, sep = "_"), paste("PrevDemMes", 0:23, sep = "_"),
    paste("PrevEol", 0:23, sep = "_"), "ResHidro", "NuclDisp", "OMIP")

  pr = c(unlist(compute(nn, data.nou, rep = 12)$net.result))
  inp = unlist(PreusI)
  obs = unlist(Preus0)
  plot(obs, type = "o")
  lines(pr, col = 2, type = "o")
}

```

```

lines(inp, col = 3, type = "o")

mapehores = abs(obs - pr)/obs
(MAPERReal[k] = mean(mapehores))
(MAPECorregit[k] = mean(mapehores[mapehores < 1]))
}

data.frame(weekdays(dates(dataeval)), dates(dataeval), MAPERReal, MAPECorregit)

```